

Diagnose lernabhängiger Veränderung mentaler Modelle

Entwicklung der SMD-Technologie als methodologisches Verfahren zur relationalen, strukturellen und semantischen Analyse individueller Modellkonstruktionen

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

der Wirtschafts- und Verhaltenswissenschaftlichen Fakultät der Albert-
Ludwigs-Universität Freiburg im Breisgau

vorgelegt von

Dirk Ifenthaler

aus Müllheim / Baden

Sommersemester 2006

Erstgutachter

Prof. Dr. Norbert M. Seel

Zweitgutachter

Prof. Dr. Thomas Eckert

Dekan der Wirtschafts- und
Verhaltenswissenschaftlichen Fakultät

Prof. Dr. Dr. Jürgen Bengel

Datum des Promotionsbeschlusses

24. Oktober 2006

Für meine Eltern
Rita und Karlfrieder

VORWORT

Was sind mentale Modelle? Warum konstruieren Individuen mentale Modelle? Warum sind mentale Modelle subjektiv plausibel? Wie werden mentale Modelle repräsentiert? Können Individuen mentale Modelle externalisieren? Verändern sich mentale Modelle in Lernprozessen? Können pädagogische Interventionen die mentale Modellbildung beeinflussen? Ist eine Diagnose mentaler Modelle möglich? Welche Diagnoseinstrumente erfassen das Konstrukt der mentalen Modelle?

Die Forschung zu mentalen Modellen weist eine umfangreiche Tradition im Bereich der Kognitions- und Lernpsychologie auf. Viele der eingangs gestellten Fragen werden in unterschiedlichen Kontexten beleuchtet und kontrovers diskutiert. Anknüpfend daran, möchte sich die vorliegende Arbeit aus methodologischer Sicht an dieser Diskussion beteiligen. Die zentrale Frage nach einer objektiven, reliablen und validen Diagnose lernabhängiger Veränderung mentaler Modelle setzt neben der theoretischen Auseinandersetzung mit dem Konstrukt der mentalen Modelle die Entwicklung eines Diagnoseinstrumentes voraus. Das neu entwickelte Verfahren der *SMD-Technologie* ermöglicht unabhängig von einer Inhaltsdomäne, eine manuelle oder automatisierte Diagnose externalisierter Modelle. Dabei werden die Modelle auf drei Ebenen mit unterschiedlichem Fokus diagnostiziert.

In drei experimentellen Studien wird der zentralen Forschungsfrage nachgegangen, ob bzw. wie sich mentale Modelle lernabhängig verändern. Das längsschnittliche Untersuchungsdesign ermöglicht eine präzise Diagnose über insgesamt sieben Messzeitpunkte. Experimentelle Variationen und unterschiedliche Untersuchungsgruppen ermöglichen darüber hinaus eine Analyse pädagogischer Interventionen während des Lernprozesses.

Die Arbeit schließt mit einer Diskussion der empirischen Befunde und einem Ausblick, welcher die Grenzen der Untersuchung beleuchtet. Außerdem lassen sich aus den empirischen Befunden weitere Forschungsfelder identifizieren und aussichtsreiche Entwicklungspotentiale ableiten.

Forschungsvorhaben entfalten ihr eigentliches Potential durch den wissenschaftlichen Diskurs. So entstanden in zahlreichen Gesprächen wichtige Ideen

und Ansatzpunkte für das Forschungsdesign, die Entwicklung des Diagnoseinstrumentes und die methodologische Auseinandersetzung mit Längsschnittuntersuchungen.

Mein besonderer Dank gilt Prof. Dr. Norbert M. Seel, für die wichtigen Impulse und vielen anregenden Gespräche und Diskussionen. Das perfekte Mengenverhältnis zwischen größtmöglichem Freiraum, hervorragenden Arbeitsbedingungen und konstruktiver Kritik hat wesentlich zum Gelingen dieser Arbeit beigetragen.

Prof. Dr. J. Michael Spector möchte ich für die umfangreichen Diskussionen in Syracuse, NY und Tallahassee, FL danken. Auch die hilfreichen Ideen von Prof. Dr. Klaus-Peter Wild und Prof. Dr. Jörg Doll haben das Forschungsvorhaben einen Schritt weiter gebracht.

Mein Dank gilt weiter dem DFG-Projektteam Bettina Couné und Ulrike Hanke sowie allen studentischen Hilfskräften, für die kollegiale, professionelle und enge fachliche Zusammenarbeit. Die drei empirischen Studien dieser Arbeit hätten ohne die Bereitschaft und das Engagement vieler Schüler und Studenten nicht durchgeführt werden können. Allen gilt daher ein herzliches Dankeschön.

Allen Mitarbeiterinnen und Mitarbeitern am Institut für Erziehungswissenschaft der Albert-Ludwigs-Universität Freiburg im Breisgau danke ich für die zahlreichen Gespräche und umfangreiche Unterstützung. Ein ganz besonderer Dank gilt an dieser Stelle Pablo Pirnay-Dummer, für unendliche Fachsimpelien und tiefgründige Gespräche und die freundschaftliche Kooperation bei wissenschaftlichen Beiträgen und Konferenzen.

Schließlich gilt mein Dank Volker Schweinbenz, Kathrin Egenhofer und meiner Familie, die mich stets ermutigt und herzlich unterstützt haben.

Freiburg, im Juni 2006

In der vorliegenden Arbeit wird aus Gründen der Lesbarkeit auf die reichende Listung von fast identischen Substantiven verzichtet, wenn Gattungsbegriffe verwendet werden, denen keine geschlechtsspezifische Bedeutung (z. B. Proband) zugeordnet ist. Wenn geschlechtsspezifische Aussagen gemacht werden, so werden die entsprechenden Substantivformen (z. B. Schülerin und Schüler) verwendet. Alle wissenschaftlichen Notationen sind in Anlehnung an das APA Publication Manual realisiert (American Psychological Association, 2005).

INHALTSÜBERSICHT

VORWORT	I
INHALTSÜBERSICHT	III
INHALTSVERZEICHNIS.....	IV
1 EINLEITUNG.....	1
2 MODELLZENTRIERTES LERNEN	6
3 DIAGNOSE MENTALER MODELLE.....	23
4 ENTWICKLUNG DER SMD-TECHNOLOGIE	38
5 VERÄNDERUNGSMESSUNG ALS VERFAHREN DER EMPIRISCHEN LEHR-LERN-FORSCHUNG.....	57
6 FRAGESTELLUNGEN UND HYPOTHESEN	74
7 METHODE	80
8 ERGEBNISSE DER UNTERSUCHUNGEN	109
9 DISKUSSION DER ERGEBNISSE.....	162
10 ZUSAMMENFASSUNG UND AUSBLICK.....	174
LITERATUR	177
ABBILDUNGSVERZEICHNIS.....	196
TABELLENVERZEICHNIS	199
ANHANG	203

INHALTSVERZEICHNIS

VORWORT	I
INHALTSÜBERSICHT	III
INHALTSVERZEICHNIS.....	IV
1 EINLEITUNG.....	1
1.1 PROBLEMSTELLUNG	2
1.2 ZIELSETZUNG UND AUFBAU DER EMPIRISCHEN FORSCHUNGSARBEIT	3
2 MODELLZENTRIERTES LERNEN	6
2.1 MENTALE MODELLE.....	7
2.1.1 Schemata und mentale Modelle.....	7
2.1.2 Mentale Modelle im pädagogischen Kontext	11
2.1.3 Mentale Modelle und Problemlösen.....	16
2.2 LERNABHÄNGIGE VERÄNDERUNG MENTALER MODELLE	17
3 DIAGNOSE MENTALER MODELLE.....	23
3.1 EXTERNALISIERUNG VON MENTALEN MODELLEN	24
3.2 DIAGNOSEVERFAHREN ZUR ERFASSUNG VON WISSENSSTRUKTUREN	26
3.2.1 Verfahren des lauten Denkens.....	26
3.2.1.1 Realisierung	27
3.2.1.2 Methodologie	27
3.2.1.3 Charakteristiken des Verfahrens.....	28
3.2.2 Struktur-lege-Techniken.....	28
3.2.2.1 Realisierung	29
3.2.2.2 Methodologie	29
3.2.2.3 Charakteristiken des Verfahrens.....	30

3.2.3 <i>Concept-Mapping-Tools</i>	30
3.2.3.1 Realisierung	31
3.2.3.2 Methodologie	31
3.2.3.3 Charakteristiken des Verfahrens.....	32
3.2.4 <i>Test für Kausalmodelle</i>	32
3.2.4.1 Realisierung	33
3.2.4.2 Methodologie	33
3.2.4.3 Charakteristiken des Verfahrens.....	34
3.2.5 <i>Oberflächen-Struktur-Vergleich</i>	34
3.2.5.1 Realisierung	34
3.2.5.2 Methodologie	34
3.2.5.3 Charakteristiken des Verfahrens.....	35
3.3 KRITISCHE ANALYSE DER DISKUTierten DIAGNOSEVERFAHREN	35
4 ENTWICKLUNG DER SMD-TECHNOLOGIE	38
4.1 VORÜBERLEGUNGEN.....	39
4.2 EXTERNALISIERTE MODELLE ALS GRAPHEN	39
4.2.1 <i>Umfang</i>	41
4.2.2 <i>Verknüpfungsdichte</i>	41
4.2.3 <i>Zerklüftetheit</i>	41
4.2.4 <i>Knotenzentralität</i>	42
4.2.5 <i>Kritische Anmerkung zur Verwendung von Indizes</i>	42
4.3 KONSTRUKTION UND AUFBEREITUNG DER MODELLE.....	43
4.3.1 <i>Richtsätze zur Transformation der MODELLE in Listenform</i>	44
4.4 EBENEN DER SMD-TECHNOLOGIE	45
4.4.1 <i>Surface-Structure (Perzipierte Oberflächenstruktur)</i>	45
4.4.1.1 Analyse	46
4.4.1.2 Methodologie	46

4.4.2 Matching-Structure (<i>Abgeleitete Strukturstabilität</i>)	47
4.4.2.1 Analyse	47
4.4.2.2 Methodologie	47
4.4.3 Deep Structure (<i>Semantische Strukturcharakteristik</i>)	49
4.4.3.1 Analyse	49
4.4.3.2 Methodologie	49
4.5 ANWENDUNG DER SMD-TECHNOLOGIE	51
4.5.1 Manuelle Analyse	52
4.5.2 Automatisierte Analyse	53
4.6 ZUSAMMENFASSUNG DES DIAGNOSEVERFAHRENS	56
5 VERÄNDERUNGSMESSUNG ALS VERFAHREN DER EMPIRISCHEN	
LEHR-LERN-FORSCHUNG	57
5.1 BEGRIFFSBESTIMMUNGEN	58
5.1.1 Veränderungsmessung	58
5.1.2 Variabilität und Veränderung	59
5.1.3 Messbare Größe der Veränderung	60
5.1.4 Interindividuelle Unterschiede in intraindividuellen Veränderungen	61
5.2 HISTORISCHE SYNOPSE DER VERÄNDERUNGSMESSUNG	61
5.3 SPEZIELLE PROBLEME DER VERÄNDERUNGSMESSUNG	63
5.3.1 Regressionseffekt	64
5.3.2 Reliabilitäts-Validitäts-Dilemma	66
5.3.3 Messbedeutungsproblem	67
5.3.4 Dilemmata einer Veränderungsmessung mentaler Modelle	68
5.4 TESTTHEORETISCHE MODELLE DER VERÄNDERUNGSMESSUNG	69
5.5 MODELLIERUNG VON VERÄNDERUNG MIT HLM	71
6 FRAGESTELLUNGEN UND HYPOTHESEN	74
6.1 DIAGNOSE LERNABHÄNGIGER VERÄNDERUNGEN	74

6.2 PÄDAGOGISCHE INTERVENTIONEN	76
6.3 EFFEKTE VON LERNERVORAUSETZUNGEN AUF DIE KONSTRUKTION MENTALER MODELLE	78
7 METHODE	80
7.1 ANORDNUNG DER EXPERIMENTELLEN STUDIEN	81
7.1.1 <i>Realisierung der Lernumgebung</i>	81
7.1.2 <i>Untersuchungsplan</i>	84
7.1.2.1 Experimentelle Studie 1.....	86
7.1.2.2 Experimentelle Studie 2.....	88
7.1.2.3 Experimentelle Studie 3.....	89
7.1.3 <i>Variablen der Untersuchung</i>	91
7.1.3.1 Experimentelle Variablen	92
7.1.3.2 Personenvariablen	92
7.1.3.2.1 Bereichsspezifisches Vorwissen	92
7.1.3.2.2 Allgemeine kognitive Fähigkeiten	92
7.1.3.2.3 Lernstrategien.....	92
7.1.3.2.4 Weitere Variablen	93
7.1.3.3 Abhängige Variablen	93
7.1.3.3.1 Surface-Structure Index θ	94
7.1.3.3.2 Matching-Structure Index μ	94
7.1.3.3.3 Deep-Structure Index δ	94
7.2 STICHPROBEN DER DREI EXPERIMENTELLEN STUDIEN	94
7.2.1 <i>Stichprobe der Studie 1</i>	94
7.2.2 <i>Stichprobe der Studie 2</i>	95
7.2.3 <i>Stichprobe der Studie 3</i>	95
7.3 INSTRUMENTE.....	96
7.3.1 <i>Allgemeiner Fragebogen</i>	96

7.3.2 Deklarativer Wissenstest	96
7.3.3 Berliner Intelligenz-Struktur-Test.....	97
7.3.4 LIST-Test.....	97
7.3.5 SMD-Technologie.....	97
7.3.6 Oberflächen-Struktur-Vergleich	99
7.4 METHODOLOGISCHE ANMERKUNGEN	99
7.4.1 Kritische Analyse der Bedingungsverwirklichung	100
7.4.2 Objektivität.....	101
7.4.3 Reliabilität.....	102
7.4.4 Validität.....	103
7.5 AUSWERTUNGSVERFAHREN	106
8 ERGEBNISSE DER UNTERSUCHUNGEN	109
8.1 ERGEBNISSE DER ERSTEN STUDIE	109
8.1.1 Vorbetrachtung der Untersuchungsstichprobe	109
8.1.2 Deskription der Ergebnisse	110
8.1.2.1 Diagnose lernabhängiger Veränderung	111
8.1.2.2 Pädagogische Interventionen.....	115
8.1.2.3 Effekte von Lernervoraussetzungen.....	119
8.1.3 Hypothesenprüfende Darstellung.....	120
8.2 ERGEBNISSE DER ZWEITEN STUDIE	127
8.2.1 Vorbetrachtung der Untersuchungsstichprobe	127
8.2.2 Deskription der Ergebnisse	128
8.2.2.1 Diagnose lernabhängiger Veränderung	128
8.2.2.2 Pädagogische Interventionen.....	132
8.2.2.3 Effekte von Lernervoraussetzungen.....	136
8.2.3 Hypothesenprüfende Darstellung.....	138
8.3 ERGEBNISSE DER DRITTEN STUDIE	144

8.3.1	<i>Vorbetrachtung der Untersuchungsstichprobe</i>	144
8.3.2	<i>Deskription der Ergebnisse</i>	145
8.3.2.1	Diagnose lernabhängiger Veränderung	145
8.3.2.2	Pädagogische Interventionen	149
8.3.2.3	Effekte von Lernervoraussetzungen	153
8.3.3	<i>Hypothesenprüfende Darstellung</i>	155
8.4	ZUSAMMENFASSUNG DER ERGEBNISSE	160
9	DISKUSSION DER ERGEBNISSE	162
9.1	DIAGNOSEINSTRUMENTE	162
9.1.1	<i>Oberflächen-Struktur-Vergleich</i>	163
9.1.2	<i>SMD-Technologie</i>	164
9.2	ERGEBNISSE DER EXPERIMENTELLEN STUDIEN	166
9.2.1	<i>Lernabhängige Veränderung der Modelle</i>	167
9.2.2	<i>Pädagogische Interventionen</i>	170
9.2.3	<i>Effekte von Lernervoraussetzungen</i>	171
9.3	SYNTHESE DER EMPIRISCHEN FORSCHUNGSARBEIT	172
10	ZUSAMMENFASSUNG UND AUSBLICK	174
	LITERATUR	177
	ABBILDUNGSVERZEICHNIS	196
	TABELLENVERZEICHNIS	199
	ANHANG	203

1 EINLEITUNG

„The process of model construction reveals how human memory [...] helps to create the future.“ (Holyoak, 1984, p. 217)

Die Theorie der mentalen Modelle hat in vielen wissenschaftlichen Disziplinen großes Interesse geweckt. Mittels des hypothetischen Konstrukts der mentalen Modelle wird versucht, menschliches Verhalten, Handeln und Lernen in unterschiedlichen Situationen zu erklären (vgl. Anzai & Yokoyama, 1984; Mayer, 1989; Seel & Dinter, 1995; Al-Diban, 2002). Im Gegensatz dazu, reihen sich die kritischen Auseinandersetzungen mit der Theorie der mentalen Modelle an. So wird der theoretischen Diskussion über mentale Modelle eine begriffliche Inkonsistenz vorgeworfen (vgl. Rips, 1987; Wilson & Rutherford, 1989, Johnson-Laird, 1989). Weitreichender ist die Kritik von Rouse & Morris (1986), die neben der terminologischen Unbeständigkeit die Aussagekraft empirischer Forschungsarbeiten zu mentalen Modellen in Frage stellen. Soll jedoch eine Theorie, wie die der mentalen Modelle, Erkenntnisse bezüglich komplexer Lehr- und Lernprozesse liefern, muss diese aus *theoretischer* und *empirischer Sicht* umfassend diskutiert, untersucht und analysiert werden. Darüber hinaus postuliert Scandura (1988), dass eine umfassende Lehr-Lern-Theorie *lernabhängige Veränderungen* von Wissenskonstruktionen berücksichtigen müsse.

Ausgangspunkt für die theoretische Diskussion mentaler Modelle in dieser Arbeit ist die Publikation von Seel (1991). In *Weltwissen und mentale Modelle* werden verschiedene theoretische Ansätze zu mentalen Modellen zusammengeführt, um ein einheitliches Verständnis ihres Zustandekommens und ihrer Funktion für die menschliche Informationsverarbeitung zu begründen. Seel (1991) weißt jedoch keine eigene empirische Untersuchung bezüglich mentaler Modelle im pädagogischen Kontext auf. Die vorliegende Arbeit schließt sich mit einer umfassenden empirischen Untersuchung des Konstrukts der mentalen Modelle daran an. Die Problemstellung konzentriert sich dabei auf die *pädagogische Diagnose lernabhängiger Veränderungen mentaler Modelle*, welche zunächst expliziert wird (Kapitel 1.1). Danach werden die Zielsetzung und der Aufbau der vorliegenden empirischen Forschungsarbeit hergeleitet (Kapitel 1.2).

1.1 PROBLEMSTELLUNG

Die Theorie der mentalen Modelle geht davon aus, dass Personen ihr Wissen in Form von Schemata und mentalen Modellen organisieren. Mentale Modelle haben nach Seel (2003a) die Funktionen der Vereinfachung, Veranschaulichung, Analogiebildung und Simulation.

Ein mentales Modell wird mit der Zielsetzung konstruiert, ein bestimmtes Phänomen zu erklären. Dabei wird das zu erklärende Phänomen mit dem bereits bestehenden Wissen einer Person in Beziehung gesetzt. Die Konstruktion mentaler Modelle ist also von bereits vorhandenem Wissen abhängig, was zur Folge hat, dass sich mentale Modelle von Novizen und Experten in identischen Gegenstandsbereichen deutlich unterscheiden (vgl. Seel, 1991). Diese Unterschiede können sowohl von inhaltlicher und struktureller Art sein, mentale Modelle können aber auch bezüglich deren semiotischer Repräsentation voneinander abgegrenzt werden.

Aus pädagogischer Sicht wird die Entwicklung mentaler Modelle von Novizen- hin zu Expertenmodellen als eine lernabhängige Veränderung zwischen Präkonzeptionen und kausalen Erklärungsmodellen verstanden (vgl. Snow, 1990). Die Entwicklung und der effektive Einsatz modellzentrierter Instruktionen und Lernumgebungen sind folglich erst dann möglich, wenn dynamische Veränderungen mentaler Modelle präzise diagnostiziert und aus pädagogischer Sicht interpretiert werden. Mentale Modelle verschließen sich jedoch der unmittelbaren Beobachtung und können daher nur aus beobachtbaren Daten erschlossen werden. Ein Zugang für die pädagogische Forschung ist das Verfahren der Externalisierung von mentalen Modellen. Hierfür kommunizieren modellkonstruierende Personen ihre mentalen Modelle unter Verwendung geeigneter Symbole und Zeichensysteme als externalisierte Modelle.

Im Zentrum dieser Arbeit steht folglich die Frage, ob die speziell entwickelte *SMD-Technologie* geeignet ist, *lernabhängige Veränderungen individueller Modellkonstruktionen auf relationalen, strukturellen und semantischen Ebenen zu diagnostizieren*. Darüber hinaus gilt es die Objektivität, Reliabilität und Validität des neu entwickelten Diagnoseverfahrens der *SMD-Technologie* zu überprüfen.

1.2 ZIELSETZUNG UND AUFBAU DER EMPIRISCHEN FORSCHUNGSARBEIT

Kerngedanke der vorliegenden Forschungsarbeit ist folgendes Zitat von Simon & Kaplan (1989, p. 20):

„Because good science consists of asking interesting questions that gives hope of being answered, good methodology is essential to good science.”

Ausgehend von der umfangreichen Theorie der mentalen Modelle, wird im Rahmen der vorliegenden empirischen Forschungsarbeit aus methodologischer Sicht der Frage nach einer *ökonomischen, präzisen, objektiven, reliablen und validen Diagnose lernabhängiger Veränderungen mentaler Modelle* nachgegangen. Die folgenden zehn Kapitel weisen dabei unterschiedliche Ziel- und Schwerpunktsetzungen auf.

Nach dieser Einleitung erfolgt in **Kapitel 2** eine Verortung des theoretischen Hintergrunds der mentalen Modelle. Das Kapitel zum modellzentrierten Lernen greift dabei wesentliche Eigenschaften und Funktionen mentaler Modelle nach Seel (1991) auf. Weiter erfolgt eine Positionierung des kognitions- und lernpsychologische Verständnisses von Schemata und mentalen Modellen. Eine Abgrenzung der hypothetischen Konstrukte Schemata und mentale Modelle scheint an dieser Stelle notwendig, da die im Rahmen der empirischen Untersuchung entwickelte multimediale Lernumgebung und deren inhaltliche Fragestellung die Konstruktion mentaler Modelle voraussetzt. Aus pädagogischer Sicht wird in diesem zweiten Kapitel diskutiert, ob durch instruktionale Maßnahmen gezielt auf die Konstruktion mentaler Modelle Einfluss genommen werden kann. Die Konstruktion, Modifikation und Reorganisation von mentalen Modellen erfolgt nach Seel et al. (2006) in komplexen Problemlösesituationen, was die weiteren Ausführungen dieses Kapitel verdeutlichen werden. Schließlich wird die Annahme von Johnson-Laird (1989) aufgegriffen, nach der sich die lernabhängige Veränderung mentaler Modelle von *Novizenmodellen* hin zu *Expertenmodellen* vollzieht.

Um die lernabhängige Veränderung mentaler Modelle zu diagnostizieren, sind auf Gütekriterien geprüfte Instrumente unumgänglich. **Kapitel 3** greift vor dem Hintergrund einer präzisen Diagnose lernabhängiger Veränderungen

mentaler Modelle zunächst das Verhältnis zwischen mentalen (internen) Modellen und externalisierten Modellen auf. Dabei wird verdeutlicht, dass nicht ein mentales Modell, sondern nur dessen externalisierte Form untersucht werden kann (vgl. Al-Diban, 2002). Aus methodologischer Sicht werden in Folge dessen verschiedene Diagnoseverfahren hinsichtlich deren Eignung für eine Diagnose lernabhängiger Veränderungen externalisierter Modelle begutachtet und kritisch analysiert.

Das im Rahmen dieser empirischen Forschungsarbeit entwickelte Diagnoseverfahren der *SMD-Technologie* wird in **Kapitel 4** ausführlich vorgestellt. Neben umfassenden Vorüberlegungen in Bezug auf die theoretischen und methodologischen Anforderungen eines solchen Diagnoseverfahrens, werden externalisierte Modelle in Anlehnung an die Graphentheorie definiert. Die *SMD-Technologie* ermöglicht unabhängig von einer Inhaltsdomäne eine Diagnose externalisierter Modelle auf drei Ebenen: Die *relationale, strukturelle* und *semantisch-strukturelle Analyse* ermöglicht eine umfassende Diagnose der zu verschiedenen Messzeitpunkten der Lernphase erhobenen Modelle. Die technologischen Ausführungen stellen neben der manuellen Anwendung die *SMD-Technologie* als automatisiertes Verfahren vor, welches eine hohe Durchführungsökonomie versichert.

Eine empirische Forschungsarbeit, die sich mit einer Diagnose lernabhängiger Veränderungen mentaler Modelle über mehrere Messzeitpunkte beschäftigt, muss sich aus methodologischer Sicht mit den Problemen einer Veränderungsmessung auseinandersetzen. In **Kapitel 5** erfolgt zunächst eine Einordnung und Begriffsbestimmung der in Bezug auf die Veränderungsmessung verwendeten Terminologie. Ausgehend von der Publikation von Harris (1963) reihen sich eine Vielzahl von Veröffentlichungen an, die in einer historischen Synopse der Veränderungsmessung zusammengefasst werden. Neben den speziellen Problemen einer Veränderungsmessung werden verschiedene Testmodelle für die Modellierung von Veränderungen vorgestellt und diskutiert. Schließlich wird die *Hierarchisch Lineare Modellierung* (Raudenbusch & Bryk, 2002) als aussichtsreicher Ansatz für den empirischen Teil dieser Arbeit dargestellt.

Die theoretischen Ausführungen zu mentalen Modellen, die Entwicklung der *SMD-Technologie* und die methodologische Diskussion über Veränderungs-

messung münden in die Fragestellungen und Hypothesen in **Kapitel 6**. Diese werden den Bereichen Diagnose lernabhängiger Veränderungen, pädagogische Interventionen und Effekte von Lernervoraussetzungen auf die Konstruktion mentaler Modelle zugeordnet.

Der empirische Teil sieht ein längsschnittliches Untersuchungsdesign vor, welches im Rahmen des DFG-Projektes „Modellbegründetes Lernen und Lehren. Multimedialen Gelegenheiten zum Nachdenken“ realisiert wurde. In **Kapitel 7** werden die Realisierung der multimedialen Lernumgebung, der Untersuchungsplan und die Variablen der Untersuchung aufgezeigt. Nach der Stichprobenbeschreibung der drei experimentellen Studien, wird die *SMD-Technologie* einer detaillierten Analyse bezüglich deren Objektivität, Reliabilität und Validität unterzogen. Schließlich folgen eine Begründung des methodischen Vorgehens und eine Kurzbetrachtung der statistischen Auswertungsverfahren.

Die Ergebnisdarstellung in **Kapitel 8** folgt der chronologischen Reihenfolge der drei experimentellen Studien. Diese gliedert sich jeweils in eine *Vorbetrachtung der Untersuchungsstichprobe*, einen *deskriptiven* und einen *hypothesenprüfenden Teil*. Für jede der drei experimentellen Studien konnte eine lernabhängige Veränderung der Modelle festgestellt werden. Darüber hinaus zeigen die Ergebnisse die Auswirkungen der pädagogischen Interventionen und die Effekte der Lernervoraussetzungen auf die Konstruktionen der Modelle.

Vor dem Hintergrund der Theorie der mentalen Modelle und vorliegender empirischer Forschungsarbeiten werden die Ergebnisse der drei experimentellen Studien in **Kapitel 9** diskutiert. Außerdem werden inhaltliche und methodologische Grenzen der vorliegenden empirischen Forschungsarbeit aufgezeigt, was selbstverständlich Verbesserungsvorschläge aus methodologischer Sicht impliziert. **Kapitel 10** gibt schließlich einen Ausblick auf künftige Forschungsfragen und auf die Weiterentwicklung sowie die praktische Anwendung der *SMD-Technologie*.

2 MODELLZENTRIERTES LERNEN

„Ein mentales Modell ist eine pragmatisch bedeutsame kognitive Konstruktion, mittels derer das modellschaffende kognitive System zum Ausdruck bringt, wie es die Welt ‚sieht‘ und versteht; es wird so lange beibehalten, wie es Plausibilität in bezug auf die Welt erzeugt.“ (Seel, 1991, p. 20)

Die Konzeption des modellzentrierten Lernens weist eine umfassende Tradition in der kognitions- und lernpsychologischen Forschung auf. So wird in der Psychologie durch Craik (1943) die Idee eines internen Modells als *working model* eingeführt. Stachowiak (1973) legt eine umfassende *allgemeine Modelltheorie* vor und Bandura (1979) postuliert das psychologische Paradigma des *sozialen Lernens am Modell*. Mit den Publikationen von Johnson-Laird (1983) und Gentner & Stevens (1983) wird dem Ansatz der mentalen Modelle ein besonderer Stellenwert beigemessen. Seel (1991) führt eine umfassende Abhandlung über *Weltwissen und mentale Modelle* im deutschsprachigen Raum ein. Die Bedeutsamkeit des Forschungsansatzes zeichnet sich in weiteren zahlreichen Veröffentlichungen aus (Dinter, 1993; Dutke, 1994; Seel, 1999a; Al-Diban, 2002; Held et al., 2006). Eine kürzlich veröffentlichte *Special Issue* der Zeitschrift *Technology, Instruction, Cognition and Learning* greift den Forschungsgegenstand bezüglich *Models, Simulations and Learning in Complex Domains* auf (vgl. Spector, 2006). Darüber hinaus wurden im Rahmen der AREA¹-Konferenz 2006 fünf Symposien mit aktuellen Forschungsbeiträgen zum Themenkomplex *Using Cognitive Models to Support Teaching and Learning* organisiert (vgl. Pirnay-Dummer et al., 2006).

Das Kapitel zum modellzentrierten Lernen führt zunächst in den theoretischen Hintergrund der mentalen Modelle ein (Abschnitt 2.1). Anschließend wird der für die pädagogisch-psychologische Forschung wohl interessantesten Frage nach der lernabhängigen Veränderung mentaler Modelle nachgegangen (Abschnitt 2.2).

¹ AERA ist die **A**merican **E**ducational **R**esearch **A**ssociation. Die Jahrestagung fand vom 07. bis 11. April 2006 in San Francisco, CA, USA statt.

2.1 MENTALE MODELLE

Die Theorie der mentalen Modelle (Johnson-Laird, 1983; Gentner & Stevens, 1983; Seel, 1991) kann den Situated-Cognition-Ansätzen (vgl. Greeno, 1989) als zentrales Konzept beigeordnet werden. Mentale Modelle sind kognitive Konstruktionen, die abhängig von der jeweiligen Situation und vom semantischen Wissen einer Person ad hoc konstruiert werden. Dinter (1993, p. 58) und Al-Diban (2002, p. 26) weisen auf den „zweifachen Konstruktcharakter“ mentaler Modelle hin. Demzufolge wird der Begriff des mentalen Modells als (1) hypothetisches Konstrukt der kognitionspsychologischen Wissenschaft zur Erklärung menschlicher Informationsverarbeitung und als (2) bewusste kognitive Konstruktion basierend auf einer Interaktion von menschlicher Wahrnehmung und Gedächtnis verstanden.

Neben dem Konzept der mentalen Modelle nimmt die Schematheorie eine bedeutende Position in der Kognitionspsychologie ein. Daher soll an dieser Stelle in einem Exkurs das Konzept der Schemata beleuchtet werden, bevor auf die mentalen Modelle im pädagogischen Kontext näher eingegangen wird.

2.1.1 Schemata² und mentale Modelle

Der Schema-Begriff ist neben mentalen Modellen ein weiteres hypothetisches Konstrukt der kognitions- und lernpsychologischen Forschung, welches die Organisation von Wissen im Gedächtnis beschreibt (Mandl et al., 1988). Die Einführung des Schema-Begriffs in die Psychologie wird Bartlett (1932) zugeschrieben.

Bartlett gab Versuchspersonen einen Text zu lesen, welcher nach bestimmten Zeitabständen (15 Minuten, Tage, Monate) reproduziert werden musste. Das Interesse Bartletts lag bei diesem Experiment jedoch nicht an der exakten Behaltensleistung der Versuchspersonen, sondern wie sich die Zusammenhänge der Erinnerungen über die verschiedenen Messzeitpunkte veränderten. So konnte durch Bartlett gezeigt werden, dass zum einen Versuchspersonen den Textinhalt subjektiv veränderten und Widersprüche des Textinhaltes beseitigten. Zum anderen wurden

² Es sei an dieser Stelle auf die Darstellungen von Seel (1991), Rumelhart & Ortony (1977) und Rumelhart et al. (1986) zur weiteren Vertiefung des Schemabegriffs hingewiesen.

neue Sachverhalte durch die Versuchspersonen hinzugefügt, die in der Textvorlage nicht vorhanden waren. Folglich liegt Wissen in organisierter Form als Wissensstruktur mit typischen Zusammenhängen vor, wobei eine Reproduktion von Wissen eines konstruktiven Prozesses bedarf. Diese Wissensstrukturen werden als Schemata bezeichnet.

Der Schema-Begriff wird von Schematheoretikern (vgl. Rummelhart & Norman, 1988; Anderson, 1984; Schank & Abelson, 1977; Minsky, 1975, u.a.) benutzt, um Wissensstrukturen zu beschreiben, welche typische Zusammenhänge eines Realitätsbereiches repräsentieren. In Anlehnung an Minsky (1975) wird anstelle von Schemata von *Frames* im Sinne von Situationsschemata gesprochen. Schemata, die standardisierte Ereignisfolgen repräsentieren, werden von Schank & Abelson (1977) als *Script* definiert.

Ein Beispiel für eine standardisierte Ereignisfolge wäre etwa ein Restaurantbesuch. Betritt eine Person ein Restaurant wird in Anlehnung an Schank & Abelson (1977) ein *Script* für einen Restaurantbesuch aktiviert. Dieses *Script* repräsentiert (a) die Einrichtung des Restaurants mit Tischen, Stühlen, einer Theke, der Kasse, etc., (b) die erwartenden Ereignisse wie das Zuweisen eines Tisches, der Bestellvorgang, etc., und (c) die sozialen Interaktionen zwischen Gästen und dem Restaurantpersonal. Schank (1999) geht davon aus, dass ein solches *Script* durch die Kombination von MOPs (memory organization packets) generiert wird. *Scripts* können wiederum in höheren kognitiven Strukturen, der *plans*, *goals*, *scenes* verortet werden.

Demzufolge sind Schemata hierarchisch organisiert, was bedeutet, dass Schemata auf unterschiedlichen Abstraktionsebenen gebildet werden können. Dutke (1994, p. 26) weist darauf hin, dass „allgemeine Schemata schwerer zu ändern sind als spezifische, weil die Anpassung eines hierarchisch höheren Schemas Wirkung auf alle untergeordneten hätte“. Dies würde folglich eine Reorganisation der Wissensstrukturen erfordern. Eine Modifikation eines hierarchisch niedrigeren Schemas bedeutet jedoch nicht zwangsläufig eine Reorganisation der übergeordneten Wissensstrukturen.

Schemata repräsentieren nach Seel (2003a, p. 57) „das aufgrund vielfältiger Einzelerfahrungen mit Objekten, Personen, Situation und Handlungen erworbene *verallgemeinerbare* und *abstrakte* Wissen einer Person.“ Ein Schema aufzubauen erfordert nach dieser Definition einen längeren Lernprozess. Sobald ein Schema jedoch gebildet wurde, kann dieses unmittelbar benutzt werden, um neue Informationen in vorhandene Wissensstrukturen zu assimilieren. Diese

Annahme kann auf das Äquilibrationsprinzip Piagets (1950, 1976) zurückgeführt werden. Im Sinne Piagets erfolgt Lernen durch die korrespondierenden und komplementären Prozesse der Assimilation und Akkommodation. Assimilation bedeutet dabei die Fähigkeit, Information in bereits entwickelte bzw. vorhandene Wissensstrukturen einzuordnen.

Trifft ein erfahrener Einsatzleiter der Feuerwehr am Einsatzort mit dem Alarmstichwort „Wohnungsbrand, vermisste Person“ ein, wird dieser zunächst sein schematisches Wissen über den Einsatzablauf aktivieren. Aufgrund seines Expertenwissens, welches er in zahlreichen Übungen und Einsätzen mit ähnlicher Ausgangslage entwickelt hat, wird der Einsatzleiter die Lage erkunden, evt. Gefährdungspotentiale ausfindig machen, Rettungswege identifizieren, notwendige technische Gerätschaften einkalkulieren und Einsatzbefehle für die taktischen Einheiten zur Rettung der Person und zur Brandbekämpfung formulieren bzw. einleiten. Neue Informationen zum Einsatzablauf werden in das entsprechende Schema (i. S. kognitiver Wissensstrukturen) des Einsatzleiters assimiliert. Treten in dieser Situation keine weiteren Störgrößen auf, so wird der Einsatz nach wenigen Minuten erfolgreich beendet sein.

Ein unerfahrener Einsatzleiter, der noch nie mit einer ähnlichen Situation konfrontiert wurde, kann kein schematisches Wissen aktivieren und somit auch nicht die vielen neuen Informationen der Einsatzstelle assimilieren. Dieser Zustand wird als Assimilationswiderstand bezeichnet.

Im oben skizzierten zweiten Teil des Beispiels (unerfahrener Einsatzleiter) ist das systeminterne Gleichgewicht zwischen Assimilation und Akkommodation gestört und löst somit einen kognitiven Konflikt aus. Dieser Assimilationswiderstand gilt als Bedingung für akkommodative Aktivitäten. Akkommodation bedeutet dabei die Restrukturierung, Veränderung und Neuorganisation von Wissensstrukturen (vgl. Seel, 1991, p. 45). Sollte es einer Person nicht möglich sein, ein geeignetes Schema zur Bearbeitung einer Aufgabe bzw. Lösung eines Problems zu aktivieren, wird in Entsprechung der spezifischen Anforderungssituation ein mentales Modell ad hoc konstruiert (vgl. Abbildung 1). Ist die akkommodative Aktivität erfolgreich, bzw. ist das von der Person konstruierte mentale Modell subjektiv plausibel und trägt zur Problemlösung bei, kann das systeminterne Gleichgewicht von Assimilation und Akkommodation wiederhergestellt werden.

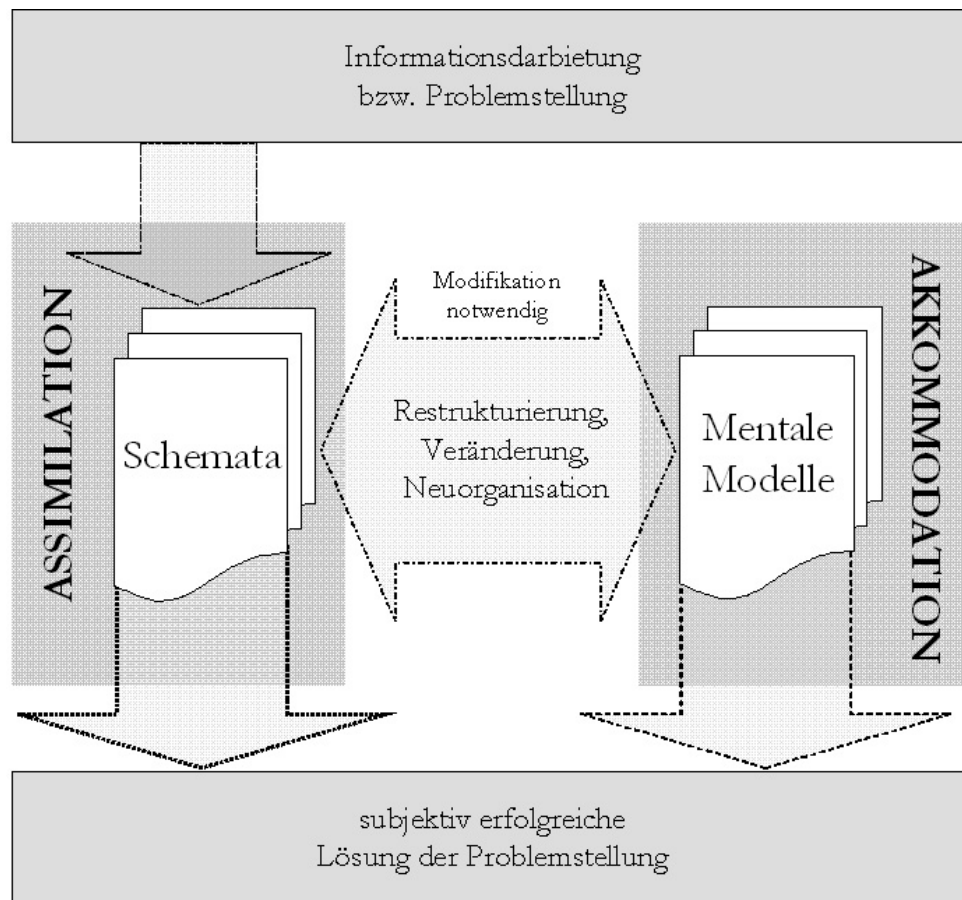


Abbildung 1: Prinzip der Äquilibration bezogen auf Schemata und mentale Modelle

Die Annahme, dass Schemata und mentale Modelle im Sinne Piagets unterschiedliche kognitive Aktivitäten erfüllen wird von verschiedenen Autoren kritisiert. Das Konstrukt der mentalen Modelle wird mit der Begründung abgelehnt, dass der Schemabegriff ausreichend sei, um die kognitiven Aktivitäten zu erfassen (vgl. Rips, 1987). Wie oben gezeigt, ist diese Annahme aus kognitionspsychologischen Gesichtspunkten nicht zweckmäßig, da Schemata und mentale Modelle unterschiedliche kognitive Aktivitäten erfüllen. Aus pädagogischer Sicht scheint es an dieser Stelle daher von Interesse zu sein, ob und wie die Konstruktion mentaler Modelle durch instruktionale Handlungen gefördert werden kann. Dieser Frage soll in den folgenden Ausführungen zu mentalen Modellen im pädagogischen Kontext nachgegangen werden.

2.1.2 Mentale Modelle im pädagogischen Kontext³

Folgt man Johnson-Lairds Darstellung (1989) bleibt zu analysieren „[...] whether there is any pedagogical advantage in providing people with models of tasks they are trying to learn.“ (p. 485). Es stellt sich also die Frage, ob durch instruktionale Maßnahmen gezielt auf die Konstruktion mentaler Modelle Einfluss genommen werden kann. Wie in den Ausführungen zu Schemata und mentalen Modellen gezeigt werden konnte, wird ein mentales Modell immer dann konstruiert, wenn neue Informationen nicht unmittelbar assimiliert werden können. Mentale Modelle werden auf der Grundlage von individuell verfügbarem Wissen einer Person ad hoc konstruiert, um eine Situation oder ein Problem zu lösen. Dabei sind mentale Modelle immer subjektiv plausibel, sind je nach Anforderungssituation inhaltlich und strukturell verschieden und werden durch verschiedene Zeichensysteme repräsentiert (vgl. Seel, 2003a, p. 258).

Der Ausgangspunkt für die theoretischen Annahmen über mentale Modelle kann auf den allgemeinen Modellbegriff zurückgeführt werden. Stachowiak legt in seiner „Allgemeinen Modelltheorie“ (1973, pp. 128) das *Abbildmerkmal*, das *Verkürzungsmerkmal* und das *pragmatische Merkmal* als die drei Hauptmerkmale des allgemeinen Modellbegriffs dar. Zum *Abbildungsmerkmal* schreibt Stachowiak (1973, p. 131):

„Modelle sind stets Modelle von etwas, nämlich Abbildungen, Repräsentationen, natürlicher oder künstlicher Originale, die selbst wieder Modelle sein können.“

Das *Verkürzungsmerkmal* von Modellen wird von Stachowiak (1973, p. 132) wie folgt definiert:

„Modelle erfassen im allgemeinen nicht alle Attribute des durch sie repräsentierten Originals, sondern nur solche, die den jeweiligen Modellerschaffern und/oder Modellbenutzern relevant erscheinen.“

Folgt man dem *pragmatischen Merkmal* nach Stachowiak, gibt es immer einen Grund für das Existieren eines Modells (ebd., pp. 132):

³ Trotz der großen Popularität des Ansatzes der mentalen Modelle seit den 1980er Jahren, konnte bei einer umfassenden Recherche mit über 600 Publikationen zu mentalen Modellen nur knapp ¼ der Titel der pädagogisch-psychologischen Forschung zugeordnet werden. Die Majorität der Publikationen befasst sich mit deduktiver Logik, Umgang mit technischen Systemen, Human-Computer-Interaction sowie Text- und Diskursverarbeitung.

„Modelle sind ihren Originalen nicht per se eindeutig zugeordnet. Sie erfüllen ihre Ersetzungsfunktion a) für bestimmte – erkennende und/oder handelnde, modellbenutzende – Subjekte, b) innerhalb bestimmter Zeitintervalle und c) unter Einschränkung auf bestimmte gedankliche oder tatsächliche Operationen.“

Ein Modell kann als eine konkrete Repräsentation von bestimmten Sachverhalten bezeichnet werden. Es bildet bestimmte Attribute eines Originals ab. Diese Attribute werden auf die notwendigen Eigenschaften reduziert, um das zu Modellierende begreiflich und durchführbar zu machen. Die Funktionen eines Modells, dessen Inhalte, strukturelle Anordnungen und dessen semiotische Repräsentationen werden durch die modellschaffende Person definiert. In verschiedenen wissenschaftlichen Disziplinen werden Modellen nachstehende vier Funktionen zugeschrieben (vgl. Seel, 2003b, p. 61):

1. Die *Vereinfachung* bezeichnet die Reduktion komplexer Phänomene auf die wesentlichen Merkmale. Vereinfachte Modelle verarbeiten daher nur die für die Klärung des Sachverhaltes notwendigen Informationen, alle irrelevanten Informationen werden dabei außer Acht gelassen.
2. Durch den Rückgriff auf Symbole, Bilder, etc. können nicht beobachtbare Phänomene als Modell veranschaulicht werden. Die *Veranschaulichung* eines Modells kann mittels unterschiedlicher Zeichensysteme erfolgen.
3. Die *Analogiebildung* bildet die Verhältnisse eines unbekannten Bereiches in Analogie zu bereits bekannten Verhältnissen ab.
4. Die *Simulation* wird als [geistiges] Experiment betrachtet, in dem mögliche Handlungen erprobt und daraus etwaige Prognosen für noch unbekannte Sachverhalte erstellt werden. Verschiedene Zustände werden simuliert, um mögliche Folgen vorhersagen zu können.

Im kognitions- und lernpsychologischen Kontext werden diese vier Modellfunktionen „didaktische Funktionen“ von mentalen Modellen gekennzeichnet (vgl. Seel, 2003a, p. 259). Die jeweilige Funktion eines mentalen Modells ist situationsabhängig und wird von einer Person nur dann benutzt, wenn diese sich in einer Problemsituation befindet und kein Schema für das entsprechende Problem verfügbar ist. Aus pädagogischer Sicht stellt eine Lernumge-

bung⁴ eine solche Problemsituation dar. Eine pädagogische Einflussnahme auf die Konstruktion mentaler Modelle kann nach Seel (1991, p. 193) auf drei Ebenen erfolgen. Auf der (1) *deklarativen Ebene* solle das relevante Weltwissen in Umfang und Qualität einer ständigen Restrukturierung unterworfen werden. Weiter sollen auf der (2) *prozeduralen Ebene* die für die mentale Modellbildung relevanten kognitiven Prozeduren verbessert werden und auf einer (3) *semiotischen Ebene* die Verwendung geeigneter Zeichensysteme für die Repräsentation von Weltwissen entwickelt werden.

Verschiedene empirische Arbeiten zeigen, dass es äußerst kompliziert aber grundsätzlich möglich ist, durch instruktionale Maßnahmen mentale Modelle zu verändern (vgl. Anzai & Yokoyama, 1984; Mayer, 1989; Seel & Dinter, 1995; Al-Diban, 2002). Die Darbietung konzeptueller Modelle⁵ (Norman, 1983) wird dabei als geeignete Methode für die Entwicklung und Konstruktion mentaler Modelle angesehen. Wird ein Lerner in einer Lernumgebung mit einem konzeptuellen Modell konfrontiert, so die Annahme, muss dieser ein mentales Modell konstruieren, um das konzeptuelle Modell zu verstehen. Da die wissenschaftlichen oder konzeptuellen Modelle eine hoch elaborierte Sichtweise eines Sachverhaltes geben, werden für Lernumgebungen oftmals instruktionale Modelle verwendet (vgl. Abbildung 2). Instruktionale Modelle sind didaktisch aufbereitete Repräsentationen konzeptueller Modelle, die in Bezug auf das Vorwissen der Lerner und die spezielle Anforderungssituation (Problemlösen im Unterricht oder in einer Alltagssituation) die Konstruktion mentaler Modelle unterstützen sollen. Die Darbietung eines instruktionalen Modells kann zu verschiedenen Zeitpunkten des Lernprozesses erfolgen.

⁴ Eine Lernumgebung wird als eine didaktische Anordnung von Lernaufgaben und Lernmaterialien in einem vorgegebenen Zeitrahmen definiert. Lernumgebungen können sowohl als *Face-to-Face-Veranstaltungen*, als reine Computerlernumgebungen oder als hybride Lernumgebungen (Computer und Face-to-Face) umgesetzt werden (vgl. Seel, 2003a).

⁵ Konzeptuelle Modelle werden als wissenschaftliche Modelle verstanden, die aus einem subjektiven mentalen Modell eines Wissenschaftlers entstanden sind und durch einen wissenschaftlichen Diskurs objektiviert wurden (vgl. Seel, 2003a).

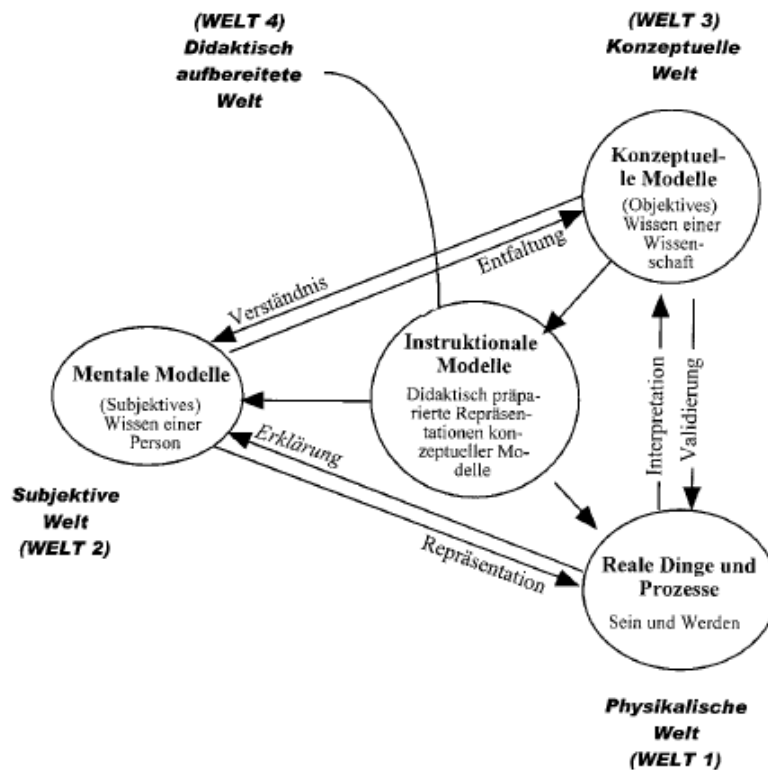


Abbildung 2: Das Verhältnis von mentalen, konzeptuellen und instruktionalen Modellen in Bezug auf die Objekt- und Ereigniswelt (Seel, 2003a, p. 261)

Mayer (1989) berichtet über experimentelle Studien, die sich mit der Darbietung von konzeptuellen Modellen aus verschiedenen Inhaltsdomänen als Lernhilfen befassen. Dabei wird zwischen „*Model Before the Lesson*“ und „*Model Within the Lesson*“ unterschieden. Folglich gibt Mayer (1989, p. 59) einen Ausblick „*How Should Conceptual Models Be Used in Instruction*“. Konzeptuelle Modelle sollen *komplett, präzise, kohärent, konzeptionell, korrekt* und *überlegt* sein. Weiter haben die Studien gezeigt, dass ein effektiver Einsatz von konzeptuellen Modellen vor oder während der Instruktion erfolgen muss. Die Darbietung nach der Instruktion wird als nicht effektiv dargestellt. Mayer (1989, p. 61) kommt zum Schluss, dass „[...] conceptual models can provide an assimilative context for students to build useful mental models“. Die experimentellen Studien von Mayer (1989) haben gezeigt, dass konzeptuelle Modelle als effektive Lernhilfen im instruktionalen Kontext eingesetzt werden können. Die Darbietung eines konzeptuellen Modells vor der Instruktion wurde auch in experimentellen Studien von Seel & Dinter (1995) als effektiv befunden.

Mentale Modelle werden von Personen in Lernprozessen mit dem Ziel konstruiert, komplexe Phänomene (vgl. Kapitel 2.1.3), reale Dinge und Prozesse

se der Welt zu erklären oder zu verstehen. Aus pädagogischer Sicht erscheint es daher sinnvoll, „'von außen' auf die Konstruktion und Entwicklung mentaler Modelle einzuwirken“ (Seel 2003a, p. 262). Instruktionale Maßnahmen können dabei auf der Basis von Sprache, Texten, Grafiken, Bildern, Diagrammen, Modellen, Animationen, Videos und Simulationen erfolgen. Aus instruktionaler Sicht sind die folgenden drei Prinzipien für die Konstruktion von mentalen Modellen (Johnson-Laird 1989, p. 487) von entscheidender Bedeutung:

„One source of mental models is observation (aided by knowledge), another is other people's explanations, and still another is our ability to construct models for ourselves either from a set of basic components or from analogous models that we already possess.“

Den drei Konstruktionsprinzipien nach Johnson-Laird (1989) folgend, werden mentale Modelle durch (a) Alltagsbeobachtungen der Welt, (b) Erklärungen anderer Personen und (c) die Fähigkeit Modelle in induktiver Weise zu konstruieren, wobei entweder auf vorhandenes Weltwissen oder bekannte analoge Modelle zurückgegriffen wird, entwickelt und aufgebaut. Diese drei Konstruktionsprinzipien lassen sich wie folgt auf instruktionale Realisierungen des modellzentrierten Lernens übertragen (vgl. Seel 2003b, p. 69):

- Selbstorganisiertes Entdeckungslernen
- Angeleitetes Entdeckungslernen
- Rezeptives Lernen, das auf die Nachahmung des Verhaltens eines Experten ausgerichtet ist

Unabhängig von den drei oben aufgeführten instruktionalen Realisierungen wird die Modellkonstruktion einer Person immer von den individuellen Vorerfahrungen und dem jeweiligen Vorwissen abhängen. Außerdem erfolgt die Konstruktion, Modifikation und Reorganisation von mentalen Modellen in komplexen Problemlösesituationen, die in speziell konzipierten Lernumgebungen realisiert werden können. Die in Kapitel 7 detailliert beschriebene multimediale Lernumgebung bietet den Lernenden „Gelegenheiten zum Nachdenken“, wobei eine komplexe Problemstellung zu einem bestimmten Phänomen gelöst werden soll. Auf die Besonderheiten von mentalen Modellen und komplexem Problemlösen soll im Folgenden eingegangen werden.

2.1.3 Mentale Modelle und Problemlösen

Es konnte bereits gezeigt werden, dass die Konstruktion eines mentalen Modells eine Problemlösesituation voraussetzt. Eine Problemlösesituation wird zunächst durch einen Ausgangs- und einen Zielzustand sowie einer Lücke zwischen diesen beiden Zuständen gekennzeichnet, welche überbrückt werden soll (vgl. Funke, 2003). Verfügt eine Person bereits über das notwendige Vorwissen und einen Regel- bzw. Handlungsplan für eine Problemlösesituation, wird das Problem durch Aktivierung von Schemata gelöst (vgl. Kapitel 2.1.1).

Im Gegensatz zu einfachen bzw. statischen Problemlösesituationen erfordern dynamische bzw. komplexe Problemlösesituationen die Konstruktion von mentalen Modellen. Ein komplexes Problem weist nach Funke (1991, 1999) folgende Merkmale auf:

- **Komplexität:** Das Problem setzt sich aus sehr vielen verschiedenen Variablen zusammen.
- **Vernetztheit:** Die Variablen sind untereinander durch wechselseitige Abhängigkeiten vernetzt.
- **Eigendynamik:** Die Problemlösesituation verändert sich zeitlich ohne die Beeinflussung des Problemlösers, was rasche Entscheidungen im Problemlöseprozess erfordert.
- **Intransparenz:** Die notwendigen Informationen für die Problemlösung sind nicht vollständig zugänglich. Folglich müssen Informationen durch den Problemlöser aktiv beschafft werden.
- **Polytelie:** Für die Problemlösung müssen multiple Kriterien und Ziele beachtet werden. Der Aufbau einer differenzierten Zielstruktur mit Regeln zur Problemlösung muss durch den Problemlöser aufgebaut werden.
- **Zeitversetzte Effekte:** Durch die dynamische Entwicklung eines Problems können Effekte während des Problemlöseprozesses auftreten, die zu Beginn noch nicht erkennbar waren.

Auf Grund der Dynamik und Vernetztheit der Variablen eines komplexen Problems ist das Aktivieren eines Schemas für das komplexe Problemlösen nicht ausreichend. Wird also eine Person mit einer komplexen Problemlösesituation konfrontiert, muss eine permanente Konstruktion, Modifikation und

Reorganisation mentaler Modelle erfolgen, um den angestrebten Zielzustand zu erreichen (vgl. Kapitel 2.1.1).

Krems (1995) zeigte in seinen Experimenten, dass sich Experten und Novizen in Situationen des komplexen Problemlösens hinsichtlich des bereichsspezifischen Wissens, deren Strategien und kognitiver Flexibilität (vgl. Spiro et al., 1988) unterscheiden. Aus seinen Experimenten in unterschiedlichen Inhaltsdomänen schließt Krems (1995, p. 214) folglich:

- Experts tend to modify their current diagnostic hypotheses more often during diagnostic reasoning than novices.
- Experts tend to use a positive test strategy to a lesser degree.
- The ability to flexibly change assumptions is likely based on case-based knowledge and does not appear to be a domain-general ability.

Bezug nehmend auf Krems (1995), ist die pädagogisch motivierte Forschung zu mentalen Modellen daher angehalten, die lernabhängige Veränderung mentaler Modelle näher zu beleuchten, um Personen bei der Modellkonstruktion eine bestmögliche Unterstützung durch instruktionale Maßnahmen bieten zu können und somit das Lösen von komplexen Problemen zu ermöglichen bzw. zu optimieren.

2.2 LERNABHÄNGIGE VERÄNDERUNG MENTALER MODELLE

„Eine Veränderung der kognitiven Struktur kann sich in jedem Teil des Lebensraumes eines Individuums, einschließlich seiner psychologischen Zukunft, der psychologischen Gegenwart oder der psychologischen Vergangenheit, ereignen; sie kann sich auf der Realitätsebene oder auf der Irrealitätsebene (Wunsch- oder Bedürfnisebene) jedes dieser Bereiche des Lebensraumes ereignen.“ (Lewin, 1942, p. 181)

Wie die Ausführungen zu mentalen Modellen im pädagogischen Kontext bereits verdeutlicht haben, ist die Konstruktion mentaler Modelle vom vorhandenen Weltwissen, flexiblen Heuristiken und der Kodierung in Form von Zeichensystemen einer Person abhängig. Selbst geringes Vorwissen einer Person beeinflusst bereits die allgemeine Informationsverarbeitung und somit auch die Konstruktion von mentalen Modellen. Wissen ist also keine statische

Einheit, vielmehr wird Wissen immer wieder neu konstruiert und permanent modifiziert (Schank, 1982). Die Dynamik und Veränderbarkeit von mentalen Modellen im pädagogischen Kontext wird von Johnson-Laird (1989, p. 485) aufgegriffen:

„What is at issue is how such models develop as an individual process from novice to expert, and whether there is any pedagogical advantage in providing people with models of tasks they are trying to learn.”

Nach Johnson-Laird (1989) vollzieht sich die lernabhängige Veränderung mentaler Modelle von *Novizenmodellen* hin zu *Expertenmodellen*. Snow (1990) stellt die Theorie der mentalen Modelle in einem größeren konzeptuellen Zusammenhang als Veränderung zwischen *Präkonzeptionen* oder *Misskonzeptionen*, die als anfängliche mentale Modelle definiert werden, hin zu *kausalen Erklärungen*, die als angestrebte Endzustände des Lernens betrachtet werden (vgl. Seel et al., 2000). Für Seel (1991, p. 7) erscheint „[...] es zweckmäßig, zwischen Experten und Laien (bzw. ‚Novizen‘) sowie zwischen ‚wissenschaftlichen‘ und ‚Alltags-Modellen‘ zu unterscheiden [...]“. Ein Experte ist nach Seel (1991, p. 7) jemand, der „über umfangreiches und detailliertes Wissen über einen bestimmten Inhaltsbereich verfügt, das er effektiv und effizient nutzen kann, um Probleme zu lösen.“ Folglich sollten erhebliche Unterschiede zwischen Experten und Novizen festzustellen sein (vgl. Al-Diban, 2002, p. 27). Die oben dargestellten Prinzipien der mentalen Modellbildung unterscheiden sich bezüglich *wissenschaftlichen* und *Alltags-Modellen* nicht. Eine Differenzierung besteht folglich nur im Umfang des individuell verfügbaren Weltwissens bzw. der „[...] Zuordnung zu unterschiedlichen Kontexten der Modellbildung [...]“ (Seel 1991, p. 7). Ein anfängliches mentales Modell wird im Folgenden als *Novizenmodell*, der angezielte Endzustand des Lernens als *Expertenmodell* definiert (vgl. Abbildung 3).

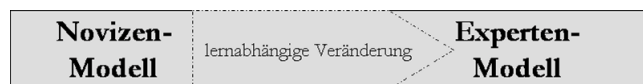


Abbildung 3: Lernabhängige Veränderung mentaler Modelle nach Johnson-Laird (1989) und Snow (1990)

Die Konstruktion mentaler Modelle im Sinne eines Novizenmodells⁶ erfolgt wie bereits dargestellt ad hoc in einer Problemlösesituation. Ein Novizenmodell „bezieht seine Plausibilität und Stimmigkeit aus seiner Übereinstimmung mit dem zugrunde gelegten bereichsspezifischen Weltwissen“ (Seel, 1991, p. 99). Novizenmodelle können auf Grund des Vorwissens einer Person sehr vage und unvollständig sein, sie beeinflussen aber dennoch die Informationsbearbeitung und somit die weitere Konstruktion mentaler Modelle (vgl. Seel & Dinter, 1995). Norman (1983, p. 8) weist auf die Unvollständigkeit anfänglicher mentaler Modelle hin:

„[...] most people’s understanding of the devices they interact with is surprisingly meager, imprecisely specified, and full of inconsistencies, gaps and idiosyncratic quirks.”

Selbst Novizenmodelle, die inhaltlich falsch sind, werden auf Grund der subjektiven Plausibilität solange von Personen beibehalten, bis diese durch bessere mentale Modelle abgelöst werden (vgl. Smith et al., 1993). Wird die subjektive Plausibilität eines mentalen Modells in Folgesituationen als erfolgreich bestätigt, „desto stabiler und fester wird es im Bewusstsein eines KS [kognitives System] verankert, wobei es häufig unerheblich zu sein scheint, ob das mentale Modell korrekt ist“ (Seel, 1991, p. 123).

Betrachtet man die *Modellveränderung* von Novizenmodellen hin zu Expertenmodellen, so wird diese als mehrstufiger Prozess der Modellbildung und –revision in einer Lern- bzw. Problemlösesituation verstanden (vgl. Penner, 2001). Dieser Prozess wird von Johnson-Laird (1983) als „fleshing-out“ und von Seel (1991) als „sukzessive Modellverständigung“ verstanden. Mit dem Prozess der Modellveränderung ist folglich eine permanente Ergänzung von Informationen verbunden, wodurch Novizenmodelle ausdifferenziert, verbessert und Expertenmodellen immer ähnlicher werden. Die Modellveränderung erfolgt auf den folgenden drei Ebenen:

⁶ Der Konstruktion von Novizenmodellen wird das *bottom-up* Prinzip zugrunde gelegt. Bedingt durch unzureichendes bereichsspezifisches Wissen werden neue Informationen Stück für Stück zu einem mentalen Modell verarbeitet. Im Gegensatz zum *bottom-up* Prinzip wird ein Experte ein mentales Modell nach dem *top-down* Prinzip konstruieren, indem bereits vorhandenes bereichsspezifisches Wissen und neue Informationen zur Modellbildung herangezogen werden (vgl. Seel & Dinter, 1995).

- Auf der *deklarativen Ebene* wird das relevante Weltwissen in Umfang und Qualität einer ständigen Restrukturierung unterworfen
- Auf der *prozeduralen Ebene* werden relevante kognitive Prozeduren und Heuristiken verbessert
- Auf der *semiotischen Ebene* werden geeignete Zeichensysteme für die Repräsentation des Weltwissens entwickelt und verfeinert

Expertenmodelle unterscheiden sich von Novizenmodellen folglich nicht nur bezüglich des bereichsspezifischen Wissens bzw. Weltwissens einer Person, sondern auch bezüglich den zur Verfügung stehenden Heuristiken und der Kompetenz der semiotischen Kodierung.

Ein Expertenmodell im Sinne eines wissenschaftlichen Modells (vgl. Abbildung 3) muss jedoch nicht immer der angezielte Endzustand eines Lernprozesses darstellen - die Veränderung mentaler Modelle hin zu einem Expertenmodell führt nicht zwangsläufig zu besserem Problemlöseverhalten. Ein verändertes mentales Modell kann zwar durch Begriffspräzisierung die externe Welt besser beschreiben, was aber zu einer gesteigerten Komplexität des mentalen Modells führt. Dies verändert wiederum das Problemlöseverhalten, da die höhere Komplexität eine ad hoc Konstruktion eines mentalen Modells einschränken wird. Ein mentales Modell kann somit zu ein und demselben Sachverhalt in unterschiedlicher Differenziertheit konstruiert werden. Larkin (1983) weist darauf hin, dass Novizenmodelle anfangs Oberflächeneigenschaften des Sachverhaltes darstellen, während Expertenmodelle außerdem strukturelle und funktionale Eigenschaften des Sachverhaltes berücksichtigen.

Um die Ausführungen zu verdeutlichen, soll das oben bereits geschilderte Beispiel aus dem Alltag eines Einsatzleiters der Feuerwehr neu aufgegriffen werden. Wiederum wird der unerfahrene Einsatzleiter alarmiert: „Technische Hilfeleistung, Schwimmbad, mehrer Personen in Gefahr“. Nach der Lageerkundung steht fest, dass die Chlorungsanlage des Schwimmbades eine Leckage aufweist und giftiges Chlorgas entweicht. Ein aufgeregter Fachmediziner versucht dem Einsatzleiter in seiner Fachsprache die chemischen und medizinischen Vorgänge einer Chlorgasintoxikation zu verdeutlichen. Der Fachmediziner externalisiert also sein Expertenmodell, um dem Einsatzleiter (Novize) die Gefahrenpotentiale (Blockade der Atmungskette, Kopfschmerzen, Angst, Atemnot, Sehstörungen, etc.) für die anwesenden Menschen im Schwimmbad zu verdeutlichen. Durch die hohe Komplexität des Expertenmodells und zusätzliche Faktoren wie Stress wird der Einsatzleiter schnell überfordert sein und somit keine fehlerfreie Entscheidung für einen reibungslosen

Einsatzablauf treffen können. Der Fachmediziner muss dem Einsatzleiter folglich die relevanten Informationen (Chlorgasdämpfe sind giftig und ätzend und werden durch die Atemwege aufgenommen, was eine Gefahr für den menschlichen Organismus darstellt) seines Expertenmodells vermitteln, um eine schnelle und fehlerfreie Einsatzentscheidung durch den Einsatzleiter zu garantieren. Das nun neu konstruierte mentale Modell des Einsatzleiters (mit den notwendigen Informationen, das komplexe Problem zu lösen), welches zwischen einem *Novizenmodell* und einem *Expertenmodell* besteht, soll im Folgenden als *Erklärungsmodell* definiert werden.

Die Abgrenzung eines *Erklärungsmodells* zu einem *Novizen-* bzw. *Expertenmodell* liegt auf der Hand. Ein *Erklärungsmodell* beinhaltet alle notwendigen Informationen, um ein Problem bezüglich des Sachverhaltes und der Anforderungssituation richtig zu lösen. Einem *Erklärungsmodell* wird dabei ein hoher Grad an Nützlichkeit beigemessen, was in Bezug auf die kognitive Leistung zu einer ergonomischen Problemlösung führt. Je nach Komplexität des Sachverhaltes und der damit verbundenen Anforderungssituation kann ein *Erklärungsmodell* einem *Novizenmodell* oder einem *Expertenmodell* sehr ähnlich sein.

DEFINITION: Ein *Erklärungsmodell* ist ein Modell, das die minimal notwendigen Informationen beinhaltet, um ein Problem zu einem bestimmten Sachverhalt in einer damit verbundenen Anforderungssituation korrekt zu lösen. Je komplexer sich der Sachverhalt darstellt, desto ähnlicher werden sich *Erklärungsmodell* und *Expertenmodell* (vgl. Abbildung 4).

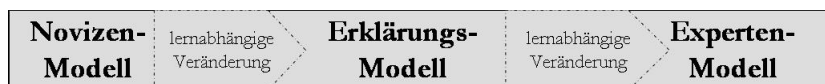


Abbildung 4: Erweiterte Darstellung der lernabhängigen Veränderung mentaler Modelle

Die lernabhängige Veränderung mentaler Modelle wird als dynamische Progression von Novizenmodellen hin zu Erklärungs- und/oder Expertenmodellen verstanden. Eine Modellveränderung bzw. ein erfolgreicher Lernprozess kann je nach Anforderungssituation auf den oben skizzierten deklarativen, prozeduralen und semiotischen Ebenen erfolgen. Die lernabhängige Veränderung mentaler Modelle kann „durch gezielte externe Einflussnahmen (z. B. durch Instruktion)“ (Seel, 1991, p. 9) manipuliert werden. Soll von außen auf die Konstruktion und Veränderung mentaler Modelle eingewirkt werden, wird das instruktional Dargebotene jedoch mit der subjektiv plausiblen Sichtweise der Person konfrontiert. Bedingt durch die subjektive Plausibilität weisen men-

tale Modelle jedoch eine hohe Veränderungsresistenz auf (vgl. McCloskey, 1983). Seel (1999a, p. 159) beschreibt dieses pädagogische Dilemma wie folgt:

„In fact, students’ preconceptions which are plausible from their subjective perspective may compete with instructionally presented conceptions that often lack perceptual evidence and seem to be inconsistent with the students’ everyday experiences. As a consequence, preconceptions are not easy to modify or to change if an individual is convinced of their plausibility.”

Verschiedene empirische Studien haben jedoch gezeigt, dass es durchaus möglich ist, von außen auf die Konstruktion und Veränderung mentaler Modelle einzuwirken (vgl. Anzai & Yokoyama, 1984; Mayer, 1989; Seel & Dinter, 1995; Al-Diban, 2002). Obwohl die oben zitierten Studien auf die Veränderung mentaler Modelle hinweisen, bleibt die Frage nach einer präzisen Diagnose lernabhängiger Veränderung mentaler Modelle offen (vgl. Ifenthaler & Seel, 2005). Die Forschung zu mentalen Modellen gleicht oft der *Black Box* des klassischen Behaviorismus, indem sie sich auf das beobachtbare und beschreibbare Verhalten vor und nach dem Lernprozess konzentriert. Dadurch werden aus pädagogischer Sicht wichtige Aspekte des Informationsverarbeitungsprozesses der mentalen Modellbildung außer Acht gelassen. Nur durch eine exakte Diagnose der lernabhängigen Veränderung mentaler Modelle zu verschiedenen Zeitpunkten während des Lernprozesses, von Novizenmodellen hin zu elaborierteren Modellen bzw. Erklärungs- und/oder Expertenmodellen, können ausreichend Anhaltspunkte aufgezeigt und optimal abgestimmte instruktionale Maßnahmen abgeleitet und durchgeführt werden. Vor dem Hintergrund einer kleinschrittigen Diagnose mentaler Modelle beleuchten die folgenden Ausführungen daher verschiedene Ansätze zur pädagogischen Diagnose mentaler Modelle.

3 DIAGNOSE MENTALER MODELLE

„[...] it is appropriate and necessary to rethink the scientific principles and philosophical assumptions that serve as the foundations of educational assessment.”
(Pellegrino et al., 2001, pp. 32)

Die Frage nach einer reliablen und validen Diagnose mentaler Modelle stellt ein zentrales Problem der mentalen Modellforschung dar (vgl. Seel, 1997). Mentale Modelle sind hypothetische Konstrukte der Wissenschaft. Diese sind nicht unmittelbar beobachtbar und werden von Personen nur in bestimmten Situationen konstruiert. Eine Diagnose mentaler Modelle setzt voraus, dass Individuen ihre internen mentalen Modelle externalisieren. Seel (1991, p. 155) weist der Externalisierung mentaler Modelle aus methodologischer Sicht eine zentrale Bedeutung bei, „[...] da ein Außenstehender nur etwas über mentale Konstruktionen eines KS [kognitives System] ‚erfahren‘ kann, wenn sie mit Hilfe sprachlicher oder andersartiger Kommunikation im Dialog mitgeteilt wird“.

Die Tatsache, dass mentale Modelle nicht direkt, sondern nur mittels geeigneter abhängiger Variablen zu erfassen sind, begrenzt die Möglichkeiten der Diagnose mentaler Modelle auf wenige Verfahren. Die im Rahmen dieser Arbeit durchgeführte Recherche konnte nur wenige Verfahren und Instrumente identifizieren, die den geforderten Gütekriterien entsprechen. So beschränkt sich die Auswahl für eine detaillierte Analyse auf das *Verfahren des Lauten Denkens*, *Struktur-Lege-Techniken*, *Concept-Mapping-Tools*, den *Test für Kausalmodelle* und den *Oberflächen-Struktur-Vergleich*.

Vor dem Hintergrund einer präzisen Diagnose mentaler Modelle wird zunächst das Verhältnis zwischen mentalen (internen) Modellen und externalisierten Modellen aufgegriffen (Abschnitt 3.1). Die im folgenden Abschnitt 3.2 dargestellten Diagnoseverfahren haben das Ziel, Wissensstrukturen bzw. mentale Modelle, die durch Personen kommuniziert werden, zu erfassen. Ob und wie die aus den Diagnoseverfahren gewonnen empirischen Daten verwertbar sind, wird im letzten Teil dieses Kapitels (Abschnitt 3.3) kritisch beleuchtet.

3.1 EXTERNALISIERUNG VON MENTALEN MODELLEN

„Die ‚Welt der Gedanken und Vorstellungen‘ bleibt dem Außenstehenden so lange verschlossen, wie sie nicht mitgeteilt wird.“ (Seel, 1991, p. 179)

Eine Person, die ein mentales Modell konstruieren möchte, muss zunächst in der Lage sein, ihr Wissen mit Hilfe geeigneter Zeichen- und Symbolsysteme zu repräsentieren (vgl. Seel, 1991, p. 153). Nach Bruner (1964) kann zwischen *enaktiven*, *ikonischen* und *symbolischen* Formaten⁷ der Wissensrepräsentation unterschieden werden. Die Art und Weise, in welchem Format Wissen repräsentiert wird, ist von der Kompetenz und der Präferenz der modellkonstruierenden Person abhängig (vgl. Seel, 2003a, p. 61). Die Kompetenz bestimmte Repräsentationsformate zu wählen ist laut Piaget (1975) u.a. auf entwicklungsbedingte Faktoren zurückzuführen. Präferenzen für bestimmte Repräsentationsformate (z. B. Veranschaulichung eines Sachverhaltes durch Schaubilder) werden in verschiedenen empirischen Studien berichtet (vgl. Gardner, 1985).

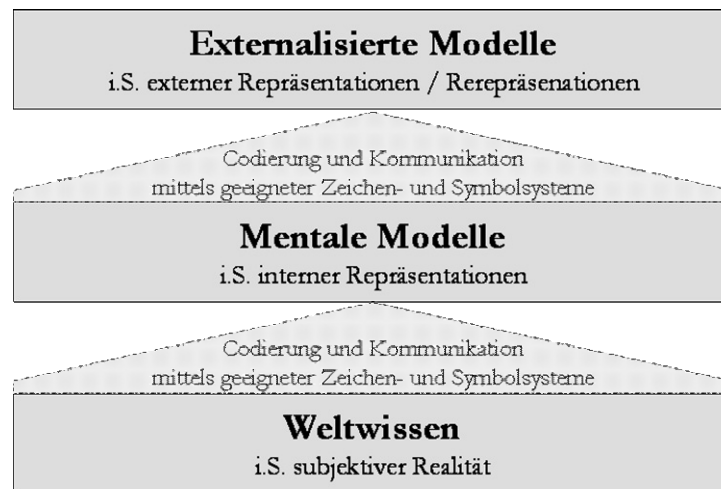


Abbildung 5: Prozess der Externalisierung mentaler Modelle

⁷ Unter *enaktiven* Formaten wird die Repräsentation in Form von Handlungen verstanden, um ein bestimmtes Ziel zu erreichen (z.B. „Ich weiß wie ich einen Autoreifen wechsle“). *Ikonische* Formate beziehen sich auf die Repräsentation in Form von Bildern (z.B. eine Handskizze über die einzelnen Schritte des Reifenwechsels an einem PKW). Die *symbolisch* Repräsentation erfolgt über durch Regeln festgelegte Symbole (z.B. Schrift und Buchstaben; mathematische Formeln).

Die Externalisierung⁸ (lat.: Nach-Außen-Verlagerung; Veräußerung) mentaler Modelle wird im Folgenden als bewusster Prozess einer Person in Form einer Codierung und Kommunikation mittels geeigneter Zeichen- und Symbolsysteme verstanden (vgl. Abbildung 5).

Eine Externalisierung mentaler Modelle ist folglich als ein zweifacher Codierungsprozess aufzufassen. Die erste Codierung wird als „interne Codierung“ bezeichnet, wobei ein mentales Modell aus dem vorhandenen Weltwissen konstruiert wird, um subjektive Plausibilität zu erzeugen. Das mentale Modell i.S. interner Wissensrepräsentationen wird mittels subjektiv geeigneter Zeichen- und Symbolsysteme repräsentiert. Die zweite Codierung erfolgt nur, wenn eine Person aufgefordert wird, ein mentales Modell mitzuteilen, oder wenn eine Person ihr mentales Modell bewusst externalisieren möchte. Dieser Prozess erfordert neben der Codierung mittels geeigneter Zeichen- und Symbolsysteme ein Format der Kommunikation, welches für andere Personen verständlich ist.

Mentale Modelle in externalisierter Form haben demzufolge die Funktion, das subjektive Weltwissen einer Person anderen oder sich selbst zu vermitteln bzw. darzustellen. Nach Seel (1991, p. 187) lassen sich die Anforderungen an externalisierte Modelle wie folgt zusammenfassen:

- Sie sollen *vollständig*,
- und *konzise* sein,
- außerdem sollen sie *kohärent* und *konkret* sein,
- und sie sollen *bedeutungshaltig* und *korrekt* sein.

Bezogen auf die Diagnose mentaler Modell lässt sich zusammenfassen, dass nicht ein mentales Modell, sondern nur dessen externalisierte Form untersucht werden kann (vgl. Al-Diban, 2002, p. 109). Hanke (2006) geht auf externe Modellbildung bei der Informationsverarbeitung und als Lernhilfen detailliert ein. Aus methodologischer Sicht bleibt jedoch offen, welche Verfahren für eine Diagnose externalisierter Modelle geeignet sind. Im folgenden Abschnitt 3.2 werden daher unterschiedliche Diagnoseverfahren hinsichtlich deren Realisierung und Methodologie diskutiert und kritisch analysiert.

⁸ Zur Repräsentation von mentalen Repräsentationen siehe u.a. Le Ny (1993).

3.2 DIAGNOSEVERFAHREN ZUR ERFASSUNG VON WISSENSSTRUKTUREN

Grundvoraussetzung für Verfahren zur Diagnose lernabhängiger Veränderung mentaler Modelle ist die Erfassung mentaler Modelle zu verschiedenen Zeitpunkten des Lernprozesses (vgl. Seel, 1999a). Weitere Grundvoraussetzung ist die Einbettung des Messverfahrens in den Kontext des komplexen Problemlösens, bzw. in Situationen in denen davon ausgegangen werden kann, dass eine Person tatsächlich ein mentales Modell konstruiert (vgl. Al-Diban, 2002).

Vor dem Hintergrund der Theorie der mentalen Modelle werden in diesem Kapitel verschiedene Diagnoseverfahren hinsichtlich ihrer Eignung zur Erfassung individueller, interindividueller und intraindividueller Wissensrepräsentationen geprüft. Dabei stehen die Merkmale und Zielsetzungen der einzelnen Diagnoseverfahren sowie deren methodologischer Zugang im Vordergrund. Die im Folgenden vorgestellten Diagnoseverfahren repräsentieren jedoch lediglich eine exemplarische Auswahl pädagogischer Standardverfahren⁹ zur Diagnose von Wissensstrukturen. Die Auswahl beschränkt sich auf das *Verfahren des Lauten Denkens*, *Struktur-Lege-Techniken*, *Concept-Mapping-Tools*, den *Test für Kausalmodelle* und den *Oberflächen-Struktur-Vergleich*.

3.2.1 Verfahren des lauten Denkens

Ausgehend vom besonderen Stellenwert der Sprache als höchst automatisierte Form der menschlichen Kommunikation, wird das *Verfahren des lauten Denkens* in der pädagogisch-psychologischen Forschung zur Erfassung von kognitiven Prozessen beim komplexen Problemlösen eingesetzt (vgl. Dörner, 1974).

⁹ Die in der Kognitionswissenschaft üblichen psychophysikalischen Methoden werden in dieser Arbeit ausgegrenzt, da diese bezüglich einer präzisen Wissensdiagnose als weniger geeignet betrachtet werden (vgl. Seel, 1997, pp. 121).

3.2.1.1 Realisierung

Das *Verfahren des lauten Denkens* (vgl. van Someren et al., 1994) setzt voraus, dass Personen während der Bearbeitung einer Problemstellung ihre kognitiven Prozesse verbalisieren. Die doppelte Anforderungssituation des gleichzeitigen Problemlösens und Verbalisierens stellt für viele Versuchspersonen eine ungewohnte Situation dar. Der Versuchsleiter ist deshalb dazu angehalten, immer wieder Detailinformationen von Versuchspersonen einzufordern, um eine möglichst präzise und kleinschrittige Analyse der kognitiven Prozesse zu ermöglichen. Bei unzureichender Verbalisierung eines Probanden werden diese durch standardisierte Instruktionen zu einer möglichst genauen Verbalisierung durch den Versuchsleiter angeleitet. Ein Training zur Verbalisierung kognitiver Prozesse kann dem eigentlichen Experiment vorgeschaltet werden.

Als wichtige Voraussetzung für das *Verfahren des lauten Denkens* gilt die Konstruktion der Problemlösesituation bzw. der Aufgabenstellung. Es kommen nur solche Aufgabenstellungen in Frage, die bezüglich der zu untersuchenden wissenschaftlichen Fragestellung spezifische Eigenschaften aufweisen.

3.2.1.2 Methodologie

Die Diagnose wird in Einzelfallanalysen realisiert. Die Durchführung erfolgt nach einem standardisierten Ablaufplan und kann in vier Phasen eingeteilt werden. Die *Einstiegsphase* dient dazu, den Probanden die ungewohnte Situation eines Experimentes vertraut zu machen. In der zweiten Phase der *Beobachtung* werden die Probanden dazu aufgefordert, Lösungsansätze für bestimmte Aufgabenstellungen zu verbalisieren. Der Versuchsleiter soll in dieser Phase Notizen zu Auffälligkeiten der Versuchsdurchführung erstellen. Außerdem soll er die Probanden zum verbalisieren animieren, Rückmeldungen bzw. Feedback geben und in ausweglosen Situationen Hilfestellungen zur Aufgabenstellung geben. Im Anschluss an die Aufgabenbearbeitung kann in einer dritten Phase ein *Interview* oder eine *Interpretation des Ton- oder Videomitschnittes* durch die Probanden erfolgen. Die vierte Phase dient der *Auswertung* und *Interpretation* der gewonnenen Daten.

Die gewonnenen Daten werden in geeigneter Form aufbereitet und in ein Kategoriensystem eingeordnet. Die Einordnung einzelner Aussagen von Pro-

banden in Kategorien erfordert ein gleichzeitiges Zuordnen der Verbaldaten zur jeweiligen Situation der Aufgabenbearbeitung während des Experimentes. Die Auswertung der gewonnenen Daten erfolgt in einer Vielfalt von Verfahren wie z. B. der Analyse von Häufigkeiten bestimmter Nennungen und Kategorienzugehörigkeiten. Beispiele zur Anwendung des Verfahrens des lauten Denkens finden sich u.a. bei Kluwe (1988).

3.2.1.3 Charakteristiken des Verfahrens

Das *Verfahren des lauten Denkens* setzt voraus, dass Probanden ihre kognitiven Prozesse der Aufgabenbearbeitung simultan verbalisieren. Es kann davon ausgegangen werden, dass die gewonnenen Daten nur eine Selektion bestimmter kognitiver Prozesse der Probanden darstellen. Oft werden unzureichende und oberflächliche Verbalisierungen als Hauptprobleme bei der Versuchsdurchführung genannt.

Die Komplexität einer Aufgabenstellung kann das Verbalisieren von Probanden zusätzlich beeinträchtigen. Auch die ungewohnte Untersuchungssituation mit Ton- und Videoaufnahmen muss bezüglich der internen Validität des Experimentes (vgl. Campbell & Stanley, 1966) beachtet werden. Die Erstellung von Kategoriensystemen zur Einordnung der Verbalprotokolle muss sich an der zu untersuchenden Theorie orientieren.

Die Problematik des *Verfahrens des lauten Denkens* liegt speziell in der Quantifizierung der gewonnenen Daten und der eindeutigen Zuordnung von Verbalisierungen zu bestimmten kognitiven Prozessen. Dies ist auf unzureichende Verbalisierungen in bestimmten Situationen zurückzuführen. Interindividuelle und speziell intraindividuelle Vergleiche erscheinen auf Grund unzureichender Datengrundlagen demzufolge fragwürdig.

3.2.2 Struktur-Lege-Techniken

Die *Heidelberger Struktur-Lege-Technik* (SLT) wurde als Dialogverfahren zwischen Proband bzw. Probandengruppen und Versuchsleiter entwickelt, um deren subjektive Theorien zu einem Gegenstandsbereiches verbal sowie optisch zu rekonstruieren (vgl. Scheele & Groeben, 1984, 1988).

3.2.2.1 Realisierung

Von einzelnen Probanden bzw. Probandengruppen werden subjektive Theorien unter Verwendung von Begriffen und benannten Relationen abgebildet. Subjektive Theorien werden von Scheele & Groeben (1988, p. 7) wie folgt definiert:

- „Kognitionen der Selbst- und Weltsicht,
- die im Dialog-Konsens aktualisier- und rekonstruierbar sind,
- als komplexes Aggregat mit (zumindest impliziter) Argumentationsstruktur,
- das auch die zu objektiven (wissenschaftlichen) Theorien parallelen Funktionen
- der Erklärung, Prognose, Technologie erfüllt,
- deren Akzeptierbarkeit als ‚objektive‘ Erkenntnis zu prüfen ist.“

Das *Verfahren der SLT* wird in unterschiedlichen Variationen als Instrument zur Wissensdiagnose eingesetzt (vgl. Eckert, 1998). Eine wesentliche Unterscheidung der Variationen lässt sich bezüglich der Vorgabe von Begriffen und Relationen durch den Versuchsleiter bzw. der selbstständigen Erarbeitung von Begriffen und Relationen durch den Probanden vornehmen. So überlässt Jonassen (1987) die Erarbeitung von Begriffen und Relationen den Probanden, Bonato (1990) lässt die Begriffswahl frei und gibt Relationen vor, Eckert & Niegemann (1995) geben Begriffe vor und lassen die Relationen selbständig erarbeiten und Putz-Osterloh et al. (1990) geben sowohl die Begriffe als auch die Relationen vor (vgl. Eckert, 1998, p. 66).

3.2.2.2 Methodologie

Vor dem Hintergrund des umfangreichen Regelwerkes der *Struktur-Legen-Techniken* wird vor der eigentlichen Versuchsdurchführung ein umfassendes Training mit den Probanden durchgeführt. Während der Versuchsdurchführung ist der Versuchsleiter angehalten Hilfen bezüglich des SLT-Verfahrens bereitzustellen. Die Durchführung gliedert sich in den häufigsten Fällen in zwei Phasen. Die erste Phase dient der *Erarbeitung von Begriffen* zum Gegenstandsreich des Experimentes. Dabei werden im Dialog mit dem Versuchsleiter bzw.

in einem Interview Definitionen und Begriffe meist auf vorgefertigte Papierkarten geschrieben. Die zweite Phase dient der *strukturellen Verknüpfung* der in der ersten Phase explizierten Begriffe. Die Begriffe werden durch verschiedenartige Relationen zueinander gesetzt (z. B. „A“ bewirkt „B“; „B“ ist abhängig von „C“).

Datengrundlage für empirische Analysen sind die entstandenen Abbilder¹⁰ aus Begriffen und Relationen. Die Diagnose erfolgt gemäß richtig bzw. falsch verwendeter Begriffe und Relationen bzw. der Übereinstimmung der Abbilder mit Expertenlösungen. Neben qualitativen Inhaltsanalysen der Abbilder werden graphentheoretische Indikatoren wie *Umfang*, *Zerklüftetheit*, *Dichte*, *Zentralität*, *Stabilität* u.a. (vgl. Weber, 1994) verwendet.

3.2.2.3 Charakteristiken des Verfahrens

Die langwierige Durchführungsdauer der *Struktur-lege-Technik* verlangt von den Probanden eine ausdauernde kognitive Leistungsbereitschaft und hohe Motivation. Eckert (1998, p. 69) misst dem SLT-Verfahren eine „geringe Durchführungs- und Auswertungsökonomie“ bei, was jedoch „durch den Einsatz des Computers wesentlich gesteigert werden“ könne. Abgesehen von einigen wenigen graphentheoretische Maßzahlen erweist sich eine quantitative Analyse der gewonnenen Daten mangels eindeutiger Indikatoren als nicht unproblematisch.

3.2.3 Concept-Mapping-Tools

Computerunterstützte *Concept-Mapping-Tools* wurden von anfänglichen Visualisierungstools hin zu Instrumenten der Wissensdiagnose und Wissensmodellierung entwickelt (vgl. Mandl & Fischer, 2000; Novak, 1998). Die computerunterstützte Realisierung bietet nach Mandl & Fischer (2000, p. 4) eine „höhere Flexibilität bei der Darstellung der Inhalte“. Außerdem können die erstellten Netzwerke „leichter verändert werden“ und lassen eine „Archivierung“, „Weitergabe“ und „Publikation der Ergebnisse“ zu. Mit der Weiterentwicklung

¹⁰ Die entstandenen Abbilder werden im Kontext von SLT auch als Netzwerke bzw. Begriffsnetze bezeichnet.

der Computer- und Netzwerktechnik eröffnen sich neue Anwendungsmöglichkeiten für computerbasierte *Concept-Mapping-Tools*, wie z. B. die zeit- und ortonabhängige Verbreitung von Wissensstrukturen in Unternehmen, Hochschulen und Institutionen u.v.m.

3.2.3.1 Realisierung

Vor dem Hintergrund grafischer Darstellungen von Begriffen und Relationen, ähnlich der unter 3.2.2 vorgestellten *Struktur-Lege-Technik*, werden Rückschlüsse auf kognitive Strukturen der Probanden vorgenommen. Verschiedene Verfahren der *Concept-Mapping-Tools* unterscheiden sich hinsichtlich ihrer experimentellen Realisierung. Die Concept-Maps werden entweder von Probanden selbst oder im Dialog mit einem Versuchsleiter erstellt. Eine weitere Variante schließt von multiplen Begriffspaaren und Logfiledaten der Probanden mittels mathematisch-graphentheoretischer Annahmen auf die kognitiven Strukturen der Probanden.

3.2.3.2 Methodologie

Ziel computerunterstützter *Concept-Mapping-Tools* ist die Diagnose von individuellem, strukturellem und semantischem Wissen der Probanden (vgl. Weber & Schumann, 2000, p. 160). Ein Concept-Map besteht aus einzelnen Begriffen oder Konzepten, die über bestimmte Relationen miteinander verknüpft sind. Die Diagnose der Concept-Maps erfolgt durch verschiedene auf mathematischen und graphentheoretischen Annahmen basierenden Dimensionen¹¹. Die Durchführung erfolgt vor dem Hintergrund einer vom Versuchsleiter definierten Aufgabenstellung.

Die gewonnenen Daten stehen bereits in elektronischer Form für weitere Analysen zur Verfügung. Es werden drei Varianten der Datenanalyse von Concept-Maps unterschieden (vgl. Stracke, 2004). Die *graphische Analyse* dient dem optischen Vergleich mehrerer Concept-Maps. Für eine *strukturelle Analyse* werden Concept-Maps als gerichtete, benannte Graphen aufgefasst, um Aussagen über deren Umfang, Verknüpfungsdichte und Zerklüftetheit zu machen

¹¹ Es werden häufig vier Dimension für die Diagnose verwendet: Umfang, Struktur, Inhalt, Stabilität (vgl. Weber & Schumann, 2000).

(vgl. Stracke, 2004, p. 69). Die *Korrespondenzanalyse* gibt Aufschluss über die Kongruenz verschiedener Concept-Maps unter Angabe eines sog. Korrespondenzkoeffizienten.

3.2.3.3 Charakteristiken des Verfahrens

Computerunterstützte Diagnoseverfahren mittels *Concept-Mapping-Tools* vermindern den Arbeitsaufwand der Datenanalyse im Vergleich zu herkömmlichen *Struktur-Lege-Techniken* um ein Vielfaches. Die meisten *Concept-Mapping-Tools* erweisen sich als inhaltsunabhängiges Computerwerkzeug, was deren Einsatz in unterschiedlichen Wissensdomänen ermöglicht.

Durch den Einsatz eines Computers können Probanden die angefertigten Concept-Maps während des Experimentes modifizieren, indem Begriffe und Relationen hinzugefügt, gelöscht oder in ihrer Position verändert werden. Unterschiedliche Forschergruppen befassen sich mit der Entwicklung und Erforschung von *Concept-Mapping-Tools*.

- COMASOTO (Concept Mapping Software Tool) ermöglicht die Konstruktion beliebiger Concept-Maps. Die Daten können für eine Analyse mit SPSS (SPSS, 2006) oder GRADAP (GRADAP, 1990) aufbereitet werden (vgl. Weber & Schumann, 2000).
- CCMaP überzeugt durch seine einfach zu bedienende Oberfläche und leicht zu erlernende Programmsteuerung. Die Datenauswertung erfolgt nach einem Export in SPSS (SPSS, 2006). In einer Weiterentwicklung von CCMaP wurden direkte Auswertungsmöglichkeiten integriert (vgl. Reiska, 2005).
- MaNET ist eine Weiterentwicklung der Netzwerk-Elaborierungstechnik (vgl. Eckert, 2000). Neben der Konstruktion von Concept-Maps kann im Analysemodul eine Vielzahl von Auswertungsprozeduren aufgerufen werden (vgl. MannheimResearchCompany, 2005).

3.2.4 Test für Kausalmodelle

Ausgangspunkte für den *Test für Kausalmodelle* sind die *Struktur-Lege-Technik* (Scheele & Groeben, 1984) und die *Kausaldiagrammdiagnostik* von Funke (1985, 1990). Bei der Entwicklung des Verfahrens wurden die theoretischen Annahmen der mentalen Modelle besonders berücksichtigt (vgl. Al-Diban, 2002).

3.2.4.1 Realisierung

Im *Test für Kausalmodelle* tragen Probanden Zusammenhänge eines dynamischen Systems in Form von Ursache-Wirkungs-Zusammenhängen in ein Kausaldiagramm ein. Die Struktur und Inhalte der Kausaldiagramme werden den Probanden nicht vorgegeben. Vor der eigentlichen Datenerhebung wird ein „Training der semiotischen Kompetenz“ (vgl. Al-Diban, 2002, p. 100) durchgeführt.

3.2.4.2 Methodologie

Die Kausaldiagrammdiagnostik verlangt von den Probanden, die Komponenten eines komplexen Systems zu identifizieren und diese dann durch Kausalbeziehungen (i.S. von Wenn-Dann-Relationen) zu verbinden. Die einzelnen Einträge der individuellen Kausaldiagramme können als *richtig* oder *falsch* bewertet werden. Das Maß *Güte der Kausaldiagramme GdK* wurde von Funke (1990) zur Identifikation der Qualität des strukturellen Systemwissens eingeführt:

$$GdK = \frac{r_i}{r_{\max}} - \frac{f_i}{f_{\max}} \quad [3.2.1]$$

Dabei bedeutet r_i die Anzahl richtiger Eintragungen und f_i die Anzahl falscher Eintragungen in ein Kausaldiagramm. r_{\max} entspricht der Summe der implementierten Systemwirkungen und f_{\max} den Nichtwirkungen mit $r_{\max} > 0$ und $f_{\max} > 0$.

Weiter wird die Qualität der subjektiven Kausaldiagramme mittels dem *Komplexitätsindex* K_i berechnet, welcher sich aus drei Variablen (1) *Summe perzipierter Wechselwirkungen*, (2) *Summe und Länge der Kausalkreisläufe* und (3) *Summe und Länge der Kausalketten* zusammensetzt (vgl. Schenk, 2003, p. 45).

$$K_i = \frac{n_{\text{pfeile}}}{n_{\text{pfeile}} + 2k_{\max}} \sqrt{n_{\text{wec}} + l_{\text{kreis1}}^2 + l_{\text{kreis2}}^2 + l_{\text{kreis3}}^2 + 1n_{\text{ket1}} + 2^2 n_{\text{ket2}} + 3^2 n_{\text{ket3}} + 4^2 n_{\text{ket4}} + 5^2 n_{\text{ket5}} + 6^2 n_{\text{ket6}}} \quad [3.2.2]$$

Dabei bedeutet n_{pfeile} die Summe der Pfeile, d.h. Kausalverknüpfungen und k_{\max} der höchste Wert für vorhandene Kausalkreisläufe bzw. Kausalketten. n_{wec} ist die Anzahl der Wechselwirkungen, l_{kreis} die Länge der Kausalkreisläufe und n_{ket} die Anzahl der Kausalketten.

Weitere Auswertungsmaße können bezüglich *Struktur*, *Kausaltiefe*, *Symmetrie*, *Prozessverlauf* und *Verknüpfungstiefe* berechnet werden. Die große Anzahl an verschiedenen quantitativen Maßen zur Diagnose der Kausalmodelle ermöglicht eine umfassende Datenanalyse.

3.2.4.3 Charakteristiken des Verfahrens

Der *Test für Kausalmodelle* wurde speziell für die Diagnose mentaler Modelle entwickelt (vgl. Al-Diban, 2002). Das Verfahren wurde an einer Stichprobe von 114 Probanden hinsichtlich dessen Objektivität, Reliabilität und Validität untersucht und als zufrieden stellend befunden (vgl. Al-Diban, 2002, p. 266).

Bedingt durch die zeitaufwendige Erhebung der Kausalmodelle mittels paper-and-pencil-Test und späterer manueller Kodierung der Daten sollte das Verfahren im Hinblick auf die Durchführungsökonomie weiterentwickelt werden (vgl. Al-Diban, 2002, p. 262). Dennoch eignet sich der *Test für Kausalmodelle* für eine Diagnostik mentaler Modelle, da dieses Verfahren eine prozessbegleitende Erfassung subjektiver Kausalmodelle mit dem gleichen Instrument zulässt.

3.2.5 Oberflächen-Struktur-Vergleich

Der *Oberflächen-Struktur-Vergleich* (SSC: Surface-Structure-Comparison) wurde im Rahmen dieser Forschungsarbeit als *erste Entwicklungsstufe der SMD-Technologie* entwickelt. Das Verfahren eignet sich für die Diagnose augenscheinlicher Ähnlichkeiten graphischer Modelle. Mittels eines stochastischen Modells werden Übergangswahrscheinlichkeiten zwischen Modellen berechnet (vgl. Ifenthaler & Seel, 2005).

3.2.5.1 Realisierung

Externalisierte Modelle werden von unabhängigen Gutachtern (*Model-Rater*) hinsichtlich augenscheinlicher Ähnlichkeiten diagnostiziert. Dafür werden zwei zu vergleichende Modelle (Modellpaare) mittels Videoprojektor den *Model-Ratern* für 5 Sekunden präsentiert. Auf Grund der Aussagen der *Model-Rater* können schließlich Übergangswahrscheinlichkeiten berechnet werden.

3.2.5.2 Methodologie

Voraussetzung für die Anwendung des *Oberflächen-Struktur-Vergleichs* ist, dass bereits externalisierte Modelle in graphischer Form (Knoten und Relationen) zu mehreren Messzeitpunkten vorliegen. Während der Präsentation der

jeweiligen Modellpaare erhalten die *Model-Rater* die Aufgabe, für eine ähnliche Struktur der projizierten Modellpaare „+“ und für eine ungleiche Struktur ein „-“, in einen Auswertungsbogen einzutragen (vgl. Anhang A).

Die ausgefüllten Fragebögen werden durch einen Diagnostiker ausgewertet, indem die relative Häufigkeit für alle Ratings pro Modellpaar berechnet wird (vgl. Bartholomew, 1967; Ifenthaler & Seel, 2005).

$$p(M) = \frac{n_M}{n} \quad [3.2.3]$$

Die schnell und ökonomisch gewonnen Daten des *Oberflächen-Struktur-Vergleichs* enthalten wertvolle Informationen bezüglich der lernabhängigen Veränderung individueller Modelle.

3.2.5.3 Charakteristiken des Verfahrens

Der *Oberflächen-Struktur-Vergleich* wurde als schnelles und ökonomisches Verfahren entwickelt, um augenscheinliche Ähnlichkeiten von Modellpaaren zu diagnostizieren. Eine Überprüfung des *Oberflächen-Struktur-Vergleichs* mittels Test-Retest-Reliabilität zeigte in mehreren Studien eine hohe Zuverlässigkeit (vgl. Kapitel 7). Das Verfahren eignet sich für eine rasche Analyse großer Datenmengen und eine präzise Diagnose oberflächlicher Veränderungen von Modellen über mehrere Messzeitpunkte. Die Analyse mittels des *Oberflächen-Struktur-Vergleichs* bietet jedoch keine Informationen bezüglich der Struktur und semantischen Qualität der Modelle.

3.3 KRITISCHE ANALYSE DER DISKUTIERTEN DIAGNOSEVERFAHREN

Im Hinblick auf die bereits dargestellten theoretischen Annahmen (vgl. Kapitel 2) kann eine Diagnose lernabhängiger Veränderung mentaler Modelle nur durch präzise operationalisierte Indikatoren erfolgen (vgl. Seel, 1997). D.h., dass sich die Diagnose mentaler Modelle aus methodologischer Sicht an den theoretischen Explikationen orientieren muss. Eine weitere notwendige Bedingung empirischer Untersuchungen zu mentalen Modellen sind reliable und valide Diagnoseverfahren bzw. -instrumente. Die exemplarisch diskutierten

Diagnoseverfahren *Lautes Denken*, *Struktur-Lege-Techniken*, *Concept-Mapping-Tools*, *Test für Kausalmodelle* und *Oberflächen-Struktur-Vergleich* weisen hinsichtlich der Diagnose mentaler Modelle Stärken und Schwächen auf.

Das Verfahren des *Lautes Denkens* ermöglicht eine Verbalisierung einzelner kognitiver Prozesse. Eine umfassende Diagnose komplexer Wissensstrukturen i.S. mentaler Modelle wird im Bezug auf die Verbalprotokolle bezweifelt. Außerdem stellt sich ein auf die Prozessdiagnostik fokussierendes Experiment für Probanden und Versuchsleiter als nahezu undurchführbar dar. Fraglich bleibt, ob eine Quantifizierung und Interpretation der Verbalprotokolle den oben genannten Kriterien einer empirisch-quantitativ orientierten Wissenschaft entsprechen.

Die *Struktur-Lege-Technik* erweist sich als ein umfassendes Verfahren, welches den Probanden ermöglicht, Abbilder aus einzelnen Begriffen und entsprechenden Verknüpfungen herzustellen. Die oft kritisierten „Defizite in der Durchführungs- und Auswertungsökonomie“ (Eckert, 1998, p. 74) bei paper-and-pencil-Tests, erfordert eine Weiterentwicklung des Verfahrens.

Die *Concept-Mapping-Tools* können im weitesten Sinne als Weiterentwicklung der *Struktur-Lege-Technik* verstanden werden, da Computer als Konstruktions- und Auswertungswerkzeuge verwendet werden. Dennoch stellen die unterschiedlichen Formen dieser Computerprogramme keine ausreichende Bedingung für eine lernprozessbegleitende Diagnose mentaler Modelle dar. Die Implementation eines *Concept-Mapping-Tools* in eine problemorientierte Lernumgebung scheitert zum einen an den technischen Besonderheiten der Lernumgebung oder der verwendeten Tools, zum anderen an der experimentellen Anordnung entsprechend der wissenschaftlichen Fragestellung und letztlich an der Domänenabhängigkeit der Tools.

Das speziell für die Diagnose mentaler Modelle entwickelte Verfahren *Test für Kausalmodelle* erwies sich „[...] im Sinne der klassischen Testtheorie als reliables, konsistentes und hinreichend spezifisches Verfahren zur Messung strukturellen Wissens über ein komplexes System [...]“ (Seel, 1997, p. 133). Dennoch stellt sich das Verfahren als zeitaufwendig dar und zeigt in Bezug auf die Durchführungs- und Auswertungsökonomie Mängel auf.

Mittels eines *Oberflächen-Struktur-Vergleichs* ist eine schnelle und ökonomische Diagnose von Veränderungen von Modellen im Längsschnitt möglich.

Allerdings bietet das Verfahren keine Möglichkeiten, die Modelle hinsichtlich ihrer semantischen Qualität zu überprüfen. Aus dieser Einschränkung wurde im Rahmen dieser Forschungsarbeit die *SMD-Technologie* entwickelt (vgl. Kapitel 4).

Auch Spector et al. (2006) nehmen Bezug auf komplexe und dynamische Problemstellungen, die Probanden in unterschiedlichen Inhaltsdomänen lösen. Mittels Kausalmodellen sollen die Probanden an Expertise gewinnen und ihre Fähigkeiten im komplexen Problemlösen verbessern. Das Diagnoseverfahren zum Projekt *Dynamic and Enhanced Evaluation of Problem Solving* (Spector et al., 2005) zeigte, dass die Konstruktion kausaler Modelle das individuelle und kollaborative Lernen unterstützt. Als Ausblick fordern Spector et al. (2005), dass das Diagnoseinstrument in einer umfassenden Lernumgebung mit komplexen Problemen eingesetzt werden soll, um umfangreiche Daten über Lernprozesse zu erhalten.

Das im folgenden Kapitel dargestellte Verfahren der *SMD-Technologie* berücksichtigt die diskutierten Kritikpunkte und orientiert sich eng an der Theorie der mentalen Modelle, um eine den Gütekriterien entsprechende, quantitative Analyse zu ermöglichen.

4 ENTWICKLUNG DER SMD-TECHNOLOGIE

„A model of cognition and learning, or a description of how people represent knowledge and develop competence in a subject domain, is a cornerstone of the assessment development enterprise.” (Pellegrino et al., 2001, p. 176)

Ausgehend von der Theorie der mentalen Modelle (Seel, 1991) und deren dargestellten lernabhängigen Veränderung (Kapitel 2.2) wurde im Hinblick auf eine Diagnose im Längsschnittdesign das Verfahren der *SMD-Technologie* (Surface-, Matching-, Deep-Structure) entwickelt. Die Entwicklung der *SMD-Technologie* erfolgte auf Grund umfangreicher Analysen bereits bestehender Verfahren und Instrumente. Wie in Kapitel 3.3 erwähnt, weisen viele Verfahren methodologische Schwächen auf. Darüber hinaus stellten sich die Verfahren als sehr zeitaufwendig und wenig flexibel bezüglich experimenteller Designs heraus. Als drittes gilt zu erwähnen, dass eine Diagnose von kognitiven Konstrukten nicht unmittelbar erfolgen. Voraussetzung für eine Diagnose lernabhängiger Veränderung mentaler Modelle ist demzufolge, dass Personen ihre mentalen Modelle mittels geeigneter Zeichen- und Symbolsysteme externalisieren. Diese externalisierten Modelle werden durch die *SMD-Technologie* auf drei Ebenen diagnostiziert. Diese erfassen die *relationale*, *strukturelle* und *semantisch-strukturelle Dimension* der externalisierten Modelle.

Die Entwicklung eines Diagnoseverfahrens erfordert umfassende Vorüberlegungen in Bezug auf die theoretischen und methodologischen Anforderungen (Abschnitt 4.1). Anschließend werden externalisierte Modelle in Anlehnung an die Graphentheorie definiert (Abschnitt 4.2). Die Konstruktion und Aufbereitung der externalisierten Modelle wird in Abschnitt 4.3 beleuchtet. Darauf aufbauend werden die *SMD-Technologie* und deren drei Ebenen dargestellt und diskutiert (Abschnitt 4.4). Abschnitt 4.5 beleuchtet die manuelle und automatisierte Anwendung der *SMD-Technologie*. In Abschnitt 4.6 werden die wesentlichen Merkmale des Diagnoseverfahrens zusammengefasst.

4.1 VORÜBERLEGUNGEN

Die Anforderungen an die *SMD-Technologie* als Diagnoseverfahren zur Erfassung lernabhängiger Veränderungen mentaler Modelle lassen sich von den theoretischen Merkmalen und Prinzipien mentaler Modelle ableiten (vgl. Kapitel 2 & 3).

Demzufolge ist ein Diagnoseverfahren zu entwickeln, welches:

1. In echte Problemlösesituationen eingebettet werden kann.
2. In unterschiedlichen Inhaltsdomänen eingesetzt werden kann.
3. Die Konstruktion, Modifikation und Reorganisation mentaler und externalisierter Modelle erlaubt.
4. Prozessbegleitend in einem Längsschnittdesign eingesetzt werden kann.
5. Die sukzessive Modellkonstruktion und -vervollständigung von Novizenmodellen hin zu Erklärungs- bzw. Expertenmodellen erfasst.
6. Besonderheiten von Erklärungs- und Expertenmodellen berücksichtigt.
7. Quantitative Daten für eine objektive, reliable und valide Diagnose liefert.
8. Eine methodologisch eindeutige Auswertung und Interpretation ermöglicht.

Im Hinblick auf die theoretisch begründeten Anforderungen und die in Kapitel 3 dargestellten Diagnoseverfahren zur Erfassung von Wissensstrukturen orientiert sich die *SMD-Technologie* an dem Verfahren der *Struktur-Legetechniken* (vgl. 3.2.2) und dem *Test für Kausalmodelle* (vgl. 3.2.4) sowie graphentheoretischen Überlegungen (vgl. Harary, 1974; Nenninger, 1980; Bonato, 1990; Tittmann, 2003).

4.2 EXTERNALISIERTE MODELLE ALS GRAPHEN

In Anlehnung an die Graphentheorie (vgl. Harary, 1974; Nenninger, 1980; Bonato, 1990; Tittmann, 2003) wird ein externalisiertes Modell (im Folgenden als MODELL bezeichnet) als indizierter, gerichteter GRAPH aufgefasst. Dabei gelten folgende Definitionen (vgl. Bonato, 1990, pp. 6):

1. Ein GRAPH $G = (K, L)$ besteht aus einer endlichen, nicht leeren Menge K von *Knoten* zusammen mit einer Menge L von zweielementigen Teilmengen von K . Jedes ungeordnete Paar $l = \{u, v\}$ von *Knoten*, das zu L gehört, ist eine *Kante* von G und man sagt, l verbindet u und v .
2. Ein GRAPH $G = (K, L)$ ist indiziert, wenn die *Knoten* der Knotenmenge K voneinander durch Namen unterschieden werden.
3. Ein gerichteter GRAPH oder DIGRAPH $D = (K, L)$ besteht aus einer endlichen, nicht leeren Menge K von *Knoten* zusammen mit einer Menge $L \subseteq K \times K$ von geordneten Paaren verschiedener *Knoten*. Die Elemente von L werden *gerichtete Kanten* oder auch einfach *Kanten* genannt.
4. Ein benannter, gerichteter GRAPH $G = (R, K, L)$ ist ein geordnetes Tripel bestehend aus einer endlichen, nicht leeren Menge R von *Relationen* (mögliche Kantennamen), einer endlichen, nicht leeren Menge K von *Knoten* und einer endlichen Menge L *gerichteter Kanten* zwischen je zwei *Knoten*. Jedes Element $l \in L$ ist selbst ein Tripel $l = (v_i r_m v_j)$ mit $v_i, v_j \in K$ und $r_m \in R$. Die Menge R besteht also aus den möglichen Arten von *Relationen*, mit denen je zwei *Knoten* aus K verbunden werden können. L besteht aus der Menge aller in einem bestimmten GRAPHEN tatsächlich existierender gerichteter benannter *Kanten* zwischen zwei *Knoten*. Ein benannter, indizierter GRAPH heißt im Folgenden MODELL M .

Ein MODELL M besteht folglich aus n Knoten K und n Relationen (Kanten) R . Eine Proposition P besteht in Anlehnung an Weber (1994, pp. 20) aus zwei Knoten K_i und einer Relation R_j . Einzelne Propositionen P lassen sich wiederum durch Relationen zu einem MODELL M verknüpfen (vgl. Abbildung 6).



Abbildung 6: Proposition P bestehend aus zwei Knoten K und einer Relation R

Die Graphentheorie stellt eine Vielzahl von Indizes zur Verfügung, mit denen die MODELLE und deren Merkmale und Besonderheiten diagnostiziert und beschrieben werden können (vgl. Bonato, 1990; Weber, 1994; Eckert,

1998; Stracke, 2004). Im Folgenden werden die zentralen Indizes *Umfang*, *Verknüpfungsdichte*, *Zerklüftetheit* und *Knotenzentralität* dargestellt.

4.2.1 Umfang

Als Umfang U eines MODELLS M wird die tatsächliche Anzahl von Relationen r bezeichnet. Der maximale Umfang U_{\max} bezeichnet die Summe aller möglichen Relationen r in einem MODELL M .

$$U_{\max} = \frac{r * (r - 1)}{2} \quad [4.2.1]$$

Werte des Umfang U können zwischen 0 (keine Relationen) und n liegen. Ein großer Wert des Umfangs U lässt auf ein umfangreiches Modell M schließen.

4.2.2 Verknüpfungsdichte

Die Verknüpfungsdichte V ist der Quotient aus der tatsächlichen Anzahl von Relationen r und der Anzahl insgesamt möglicher Relationen r_{\max} .

$$V = \frac{r}{r_{\max}} \quad [4.2.2]$$

Die Werte der Verknüpfungsdichte V liegen zwischen 0 und 1, wobei 0 keine Verknüpfungen und 1 eine vollständige Verknüpfung des Modells M bedeutet.

4.2.3 Zerklüftetheit

Die Zerklüftetheit Z eines Modells M bezeichnet die Anzahl der unabhängigen bzw. nicht miteinander verbundenen Teilmodelle (vgl. Abbildung 7).

Das Maß Zerklüftetheit Z kann die Werte zwischen 1 (vollständige Verknüpfung) und n (Summe der einzelnen Knoten K) annehmen. Ist die Zerklüftetheit Z klein, so kann von hoch verknüpftem Wissen des entsprechenden MODELLS M ausgegangen werden.

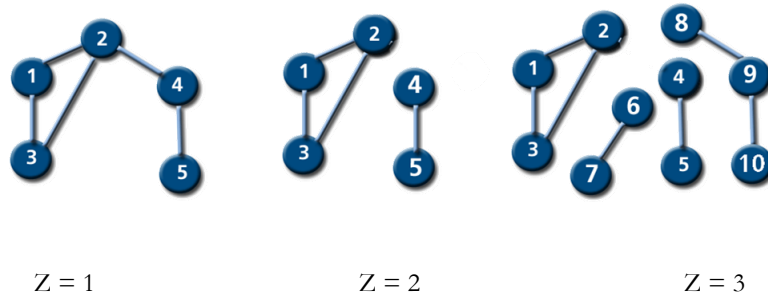


Abbildung 7: Modelle mit verschiedenen Maßen der Zerklüftetheit Z

4.2.4 Knotenzentralität

Es wird zwischen der Knotenzentralität und der Zentralität eines gesamten MODELLS unterschieden (vgl. Bonato, 1990, pp. 65). Die Knotenzentralität gibt die Gewichtung eines einzelnen Knotens innerhalb eines MODELLS an. Die Zentralität von MODELLEN kann u.a. durch die Kompaktheit, d.h. durch den Mittelwert der Distanzen über alle Knotenpaare eines MODELLS, angegeben werden.

4.2.5 Kritische Anmerkung zur Verwendung von Indizes

Es gilt an dieser Stelle anzumerken, dass graphentheoretische Indizes keine direkte Beschreibung mentaler Modelle zulassen. Die Indizes ermöglichen zwar eine mathematische Darstellung der empirisch gewonnen Daten, sie erlauben jedoch keine unmittelbaren Rückschlüsse auf die Prozesse der mentalen Modellbildung (vgl. Eckert, 1998, pp. 86). Folglich muss für eine exakte Diagnose lernabhängiger Veränderung mentaler Modelle neben der Entwicklung eines präzisen Diagnoseinstruments auch eine „systematische Konstruktvalidierung der geschilderten Indizes [...] vorgenommen“ (Eckert, 1998, p. 87) werden (vgl. Kapitel 7.4.4).

4.3 KONSTRUKTION UND AUFBEREITUNG DER MODELLE

Die Diagnose lernabhängiger Veränderung mentaler Modelle mittels der *SMD-Technologie* erfordert, dass Personen ihre mentalen Modelle in Form von Knoten und Relationen als MODELLE externalisieren. Eine nach der Theorie der mentalen Modelle entwickelte multimediale Lernumgebung stellt die Basis für drei experimentelle Studien dar (vgl. Kapitel 7). In dieser Lernumgebung konstruieren die Probanden ihre individuellen MODELLE nach einem spezifischen Notationssystem¹² in Form von Knoten und Relationen mit der Software *MS PowerPoint* (Microsoft Corporation, 2000). Die Software ermöglicht die Konstruktion, Modifikation und Reorganisation individueller MODELLE. Über verschiedene Messzeitpunkte kann somit die in der Theorie diskutierte sukzessive Modellkonstruktion und – vervollständigung von Novizen- hin zu Erklärungs- bzw. Expertenmodellen untersucht werden.

Die an definierten Messzeitpunkten (vgl. Kapitel 7) erhobenen MODELLE werden für die *SMD-Technologie* in Form von Listen aufbereitet (vgl. Abbildung 8). Neben Versuchspersonennummern und experimentellen Bedingungen werden in der Liste die Messzeitpunkte (mzp) und die einzelnen Propositionen der MODELLE erfasst. Die Propositionen bestehen jeweils aus zwei Knoten (node1, node2) und der dazugehörigen gerichteten Relation (rel).

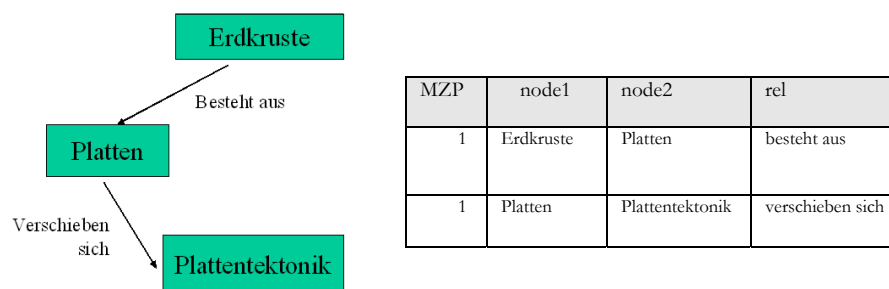


Abbildung 8: Beispiel eines Modells und dessen aufbereitete Listenform zum Messzeitpunkt 1

¹² Die Probanden wurden in einem einführenden Training in die Funktionen von *MS PowerPoint* (Microsoft Corporation, 2000) und in die Konstruktion von MODELLEN mittels Knoten und Relationen eingeführt (vgl. Kapitel 7).

Für die Transformation der erhobenen MODELLE in Listenform werden in Anlehnung an Bonato (1990) Richtsätze formuliert, die im folgenden Abschnitt dargestellt werden.

4.3.1 Richtsätze zur Transformation der MODELLE in Listenform

Um die in der empirischen Lehr-Lernforschung geforderte Auswertungsobjektivität (Bühner, 2004) zu wahren, werden folgende Richtsätze für die Transformation der MODELLE in Listenform definiert.

Satz 1: Erklärungs- und Expertenmodell

Für jede Inhaltsdomäne werden ein Erklärungs- und ein Expertenmodell erstellt (siehe Anhang B und D).

Satz 2: Modell-Lexikon

Auf der Grundlage eines jeden Erklärungs- und Expertenmodells wird ein Modell-Lexikon erstellt, welches alle Propositionen und synonyme Propositionen enthält. Das Modell-Lexikon gilt als Referenz für die dritte Ebene der *SMD-Technologie*.

Satz 3: Fehlende Knoten und Relationen

Werden in einem Modell offensichtlich fehlende Knoten oder Relationen diagnostiziert, werden diese nicht ergänzt.

Satz 4: Verknüpfung von Knoten und Relationen

Verbindet eine Relation einen Knoten diffus mit mehreren Knoten, so wird dieser nur mit dem nächstgelegenen Knoten verbunden.

Satz 5: Semantischer Inhalt der Knoten

Jeder Knoten darf laut der Anleitung zur Modellkonstruktion mittels Knoten und Relationen (vgl. Couné et al., 2003, 2004) ausschließlich ein Substantiv enthalten. Werden mehrere Substantive in einem Knoten diagnostiziert, wird das erstgenannte in die Analyse aufgenommen.

Satz 6: Semantischer Inhalt der Relationen

Jede Relation darf laut der Anleitung zur Modellkonstruktion mittels Knoten und Relationen (vgl. Couné et al., 2003, 2004) ausschließlich ein Verb enthalten. Werden mehrere Verben oder andere Wortarten in einer Relation diagnostiziert, wird das erstgenannte in die Analyse aufgenommen.

4.4 EBENEN DER SMD-TECHNOLOGIE

Die *SMD-Technologie* (**S**urface-, **M**atching-, **D**eeP-Structure) ist ein neu entwickeltes Verfahren zur Diagnose lernabhängiger Veränderung externalisierter Modelle. MODELLE werden mit der *SMD-Technologie* auf drei Ebenen diagnostiziert. Diese erfassen die *relationale* (Surface-Structure), *strukturelle* (Matching-Structure) und *semantisch-strukturelle* (Deep-Structure) Dimension. Im Folgenden soll auf die Analyse und Methodologie der drei Ebenen detaillierter eingegangen werden. Die Präzision und Diagnoseökonomie der *SMD-Technologie* kann in Analogie zu einer Pyramide dargestellt werden (vgl. Abbildung 9). Die Spitze der Pyramide stellt die Surface-Structure dar, welche eine schnelle und ökonomische Diagnose der Oberflächenstruktur individueller MODELLE ermöglicht. Im Mittelteil der Pyramide befindet sich die Matching-Structure, die auf die Strukturstabilität der MODELLE während des Lernprozesses hinweist. Die Basis der Pyramide bildet die Deep-Structure, auf welcher eine detaillierte und umfassende Analyse der semantischen Strukturcharakteristik der MODELLE erfolgt.

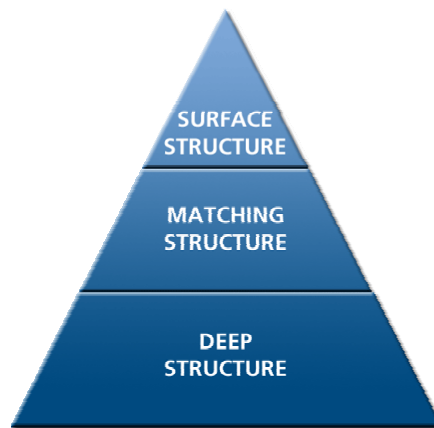


Abbildung 9: Die drei Ebenen der SMD-Technologie

4.4.1 Surface-Structure (Perzipierte Oberflächenstruktur)

Die erste Ebene der *SMD-Technologie* bildet die Surface-Structure. Auf dieser Ebene werden die MODELLE bezüglich ihrer perzipierten Oberflächenstruktur analysiert. Die relationale Analyse ermöglicht eine schnelle Diagnose

der individuellen MODELLE bezüglich der verwendeten Propositionen (vgl. Kapitel 4.2).

4.4.1.1 Analyse

Die Surface-Structure dient der qualitativen und quantitativen Analyse der individuellen MODELLE. Für die relationale Begutachtung werden die in *MS PowerPoint* (Microsoft Corporation, 2000) erstellten MODELLE hinzugezogen. Der Diagnostiker kann dabei einzelne oder mehrere MODELLE gleichzeitig einer graphisch/optischen Prüfung unterziehen.

Die *qualitative Auswertung* ermöglicht eine schnelle und ökonomische Diagnose der vorliegenden MODELLE. Soll jedoch eine Vielzahl von Modellen diagnostiziert werden, wird ein Diagnostiker aus zeitökonomischen Gründen schnell an die Grenzen der qualitativen Auswertung stoßen. Die graphisch/optische Analyse der Modelle erwies sich im Rahmen der drei experimentellen Studien dennoch als geeignete Methode, einer modellkonstruierenden Person ein individuelles Feedback bezüglich der verwendeten Knoten und Relationen bzw. des Fortschrittes im Problemlöseprozess zu geben (vgl. Kapitel 7).

Für die *quantitative Analyse* der Surface-Structure werden die individuellen Modelle in Listenform transformiert (vgl. Kapitel 4.3). Als Indikator der Surface-Structure wird die Summe aller in einem MODELL vorhandenen Propositionen eingeführt. Bei einer Vielzahl zu diagnostizierender MODELLE ermöglicht die quantitative gegenüber der qualitativen Diagnose eine schnellere und ökonomischere Auswertung. Die nachfolgenden, methodologischen Ausführungen zeigen, dass das Analyseverfahren der Surface-Structure sowohl intraindividuell als auch interindividuell anwendbar ist.

4.4.1.2 Methodologie

Als Surface-Structure Index θ eines MODELLS M wird die Summe aller vorhandenen Propositionen P bezeichnet.

$$\theta = \sum_{i=0}^n P_i \quad [4.4.1]$$

Werte des Surface-Structure Index θ können zwischen 0 (keine Proposition) und n liegen. Ein großer Wert des Surface-Structure Index θ lässt auf ein umfangreiches MODELL M schließen. Folglich kann für jedes MODELL ein Surface-Structure Index θ berechnet werden. Dieser Index kann in weitere Berechnungen einfließen. Daraus ergeben sich folgenden Optionen, die im empirischen Teil dieser Arbeit (vgl. Kapitel 6f) aufgegriffen werden:

- Intraindividueller Vergleich der MODELLE über mehrere Messzeitpunkte
- Interindividueller Vergleich der MODELLE zu gleichen und verschiedenen Messzeitpunkten
- Vergleich der MODELLE mit einem Erklärungs- bzw. Expertenmodell

4.4.2 Matching-Structure (Abgeleitete Strukturstabilität)

Die Ebene der Matching-Structure dient der strukturellen Analyse von MODELLEN. Die Struktur eines MODELLS M kann in Anlehnung an die Graphentheorie (siehe Kapitel 4.2) durch Indizes wie dem Umfang, der Zerklüftheit, der Verknüpfungsdichte, der Knotenzentralität u.a. beschrieben werden. Im Hinblick auf die Theorie der mentalen Modelle wird die *Komplexität* bzw. der *Durchmesser* eines MODELLS als aussagekräftiges Maß bestimmt, da eine Veränderung der Strukturstabilität auf die Modifikation bzw. Reorganisation eines MODELLS hinweist.

4.4.2.1 Analyse

Die quantitative Analyse der Matching-Structure setzt eine Transformation der Modelle in Listenform voraus. Als Indikator für die Konstruktion, Modifikation und Reorganisation von MODELLEN wird auf der strukturellen Ebene der *SMD-Technologie* der kürzeste Weg zwischen den am weitesten entfernten Knoten bestimmt.

4.4.2.2 Methodologie

Als Matching-Structure Index μ eines Modells M wird die Anzahl der Relationen R des kürzesten Weges zwischen den am weitesten entfernten Knoten K bezeichnet.

$$\mu = \max_{i,j} \{d(i, j)\} \quad [4.4.2]$$

Dabei ist $d(i, j)$ gleich der kleinsten Relationendistanz¹³ zwischen den Knoten i und j . Werte des Matching-Structure Index μ können zwischen 0 (keine Relationen vorhanden) und n liegen. Ein großer Wert lässt auf ein strukturell elaboriertes bzw. komplexeres MODELL schließen.

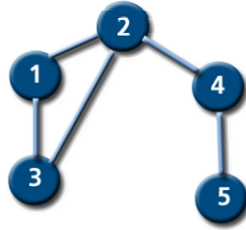


Abbildung 10: Beispiel-MODELL

Es sei ein Beispiel-MODELL M mit fünf Knoten und fünf Relationen gegeben (vgl. Abbildung 10). Für die Ermittlung des Matching-Structure Index $\mu(M)$ wird das Modell in eine Abstandsmatrix AM übertragen.

$$AM = \begin{pmatrix} 0 & 1 & 1 & 2 & 3 \\ 1 & 0 & 1 & 1 & 2 \\ 1 & 1 & 0 & 2 & 3 \\ 2 & 1 & 2 & 0 & 1 \\ 3 & 2 & 3 & 1 & 0 \end{pmatrix}$$

Der Matching-Structure Index μ ist das Maximum aller Eintragungen in der Abstandsmatrix:

$$\mu = \max_{i,j} \{d(i, j)\} = 3 \quad [4.4.3]$$

Der Matching-Structure Index für das MODELL in Abbildung 10 beträgt folglich $\mu = 3$

Jedem erhobenen MODELL M kann ein Matching-Structure Index μ zugeordnet werden, welcher zu einer umfangreichen statistischen Analyse (vgl. Kapitel 6f) herangezogen werden kann.

¹³ Als Relationendistanz wird die Anzahl der Relationen entlang eines Pfades in einem MODELL M bezeichnet.

4.4.3 Deep Structure (Semantische Strukturcharakteristik)

Die dritte Ebene der *SMD-Technologie* bildet die Deep-Structure. Auf dieser Ebene werden die MODELLE bezüglich ihrer semantischen Strukturcharakteristik diagnostiziert. Mittels eines domänenabhängigen Referenzmodells (z. B. ein Erklärungs- bzw. Expertenmodell) erfolgt eine präzise Diagnose der semantischen Struktur der individuell konstruierten MODELLE.

4.4.3.1 Analyse

Im Gegensatz zur Surface- und Matching-Structure erfolgt die Analyse der MODELLE auf der Deep-Structure nicht durch eine aus der Graphentheorie abgeleitete Indexberechnung, sondern durch eine Ähnlichkeitsbestimmung zwischen einem MODELL und dem für die Inhaltsdomäne relevanten Referenzmodell. Folglich ist eine hohe Qualität des Referenzmodells für eine umfassende Analyse auf der Deep-Structure eine notwendige und unumgängliche Bedingung. Die Ähnlichkeit der semantischen Strukturcharakteristik wird durch die Tversky-Ähnlichkeit bestimmt (Tversky, 1977).

4.4.3.2 Methodologie

Ein Ähnlichkeitsmaß drückt den Grad der Ähnlichkeit zwischen zwei Objekten durch eine Zahl aus. Ausschlaggebend für die Ähnlichkeit sind gemeinsame und unterschiedliche Merkmale der Objekte. Tversky (1977) betrachtet ein Objekt als eine Menge von Merkmalen bzw. Features. Die Bestimmung der Ähnlichkeit zweier Objekte wird durch den Vergleich der jeweiligen Merkmale bestimmt. Die Ähnlichkeitsformel berücksichtigt jedoch nicht nur die Menge der übereinstimmenden Merkmale, sondern auch die Menge der nicht vorhandenen Merkmale. Lin (1998) definiert die Ähnlichkeit mit folgenden drei Aussagen:

- The similarity between A and B is related to their commonality. The more commonality they share, the more similar they are.
- The similarity between A and B is related to the differences between them. The more differences they have, the less similar they are.
- The maximum similarity between A and B is reached when A and B are identical, no matter how much commonality they share.

Folglich ist die geringste Ähnlichkeit zwischen zwei Objekten A und B gegeben, wenn keine gemeinsamen Merkmale existieren. In diesem Fall sind die beiden Objekte komplett verschieden und der Ähnlichkeitswert ist gleich Null. Mit steigender Anzahl von gemeinsamen Merkmalen steigt der Ähnlichkeitswert. Bei einer vollständigen Übereinstimmung der Merkmale ist der Ähnlichkeitswert gleich Eins.

Die Ähnlichkeit von MODELLEN auf der Ebene der Deep-Structure wird durch das Merkmal Proposition bzw. deren semantischer Charakteristik bestimmt. Als Deep-Structure Index δ wird die Tversky-Ähnlichkeit zwischen einem MODELL M und einem Referenzmodell M_{ref} bezeichnet. Allgemein gilt:

$$\delta = \frac{f(A \cap B)}{f(A \cap B) + \alpha \cdot f(A - B) + \beta \cdot f(B - A)} \quad [4.4.4]$$

Dabei sind A und B die Mengen der Propositionen eines Modellvergleiches. Die Funktion $f(M)$ entspricht der Anzahl der Elemente in der Menge M . Durch die Parameter α und β kann die Gewichtung übereinstimmender und nicht übereinstimmender Merkmale beeinflusst werden. Werden α und β gleich gewichtet ($\alpha = \beta = 0.5$), werden sowohl die übereinstimmenden als auch die nicht übereinstimmenden Merkmale der zu vergleichenden MODELLE berücksichtigt. Werte des Deep-Structure Index δ können zwischen 0 (keine Ähnlichkeit zwischen den MODELLEN) und 1 (vollständige Übereinstimmung der MODELLE) liegen.

Sollen nur die *übereinstimmenden Merkmale* eines MODELLS in Bezug auf ein Referenzmodells diagnostiziert werden, so werden die Parameter α und β wie folgt gewichtet:

$$\alpha = \beta = \frac{f(A)}{f(B)} \cdot \frac{1}{2} \quad [4.4.5]$$

Im empirischen Teil dieser Arbeit (vgl. Kapitel 6f) wird der Deep-Structure Index δ ausschließlich mit gleich gewichteten Parametern berechnet um sowohl die *übereinstimmenden* als auch *nicht übereinstimmenden Merkmale* zu berücksichtigen. Dabei gelten für δ die Parameter $\alpha = \beta = 0.5$. Jedem erhobenen MODELL M kann folglich ein Deep-Structure Index δ zugeordnet

werden, welcher zu einer umfangreichen statistischen Analyse (vgl. Kapitel 6f) herangezogen werden kann.

4.5 ANWENDUNG DER SMD-TECHNOLOGIE

Eine umfassende Analyse individueller MODELLE mittels der *SMD-Technologie* erfordert mehrere Schritte (vgl. Abbildung 11). Als Erstes erfolgt eine präzise Messung individueller MODELLE in einem experimentellen Design (vgl. Kapitel 6f.). Die MODELLE können dabei in unterschiedlichen Repräsentationsformaten (als MODELL, Concept-Map, Aussagen in Form von Text, etc.) erfasst werden. Als zweiter Schritt erfolgt die Aufbereitung bzw. Transformation der MODELLE in Listenform (vgl. Kapitel 4.3; Lexikon siehe Anhang C und E). Die Aufbereitung bzw. Transformation erfordert die Einhaltung der in Kapitel 4.3.1 definierten Richtsätze. Als weiterer Schritt erfolgt die SMD Analyse, welche manuell (Kapitel 4.5.1) oder automatisiert (Kapitel 4.5.2) durchgeführt werden kann.

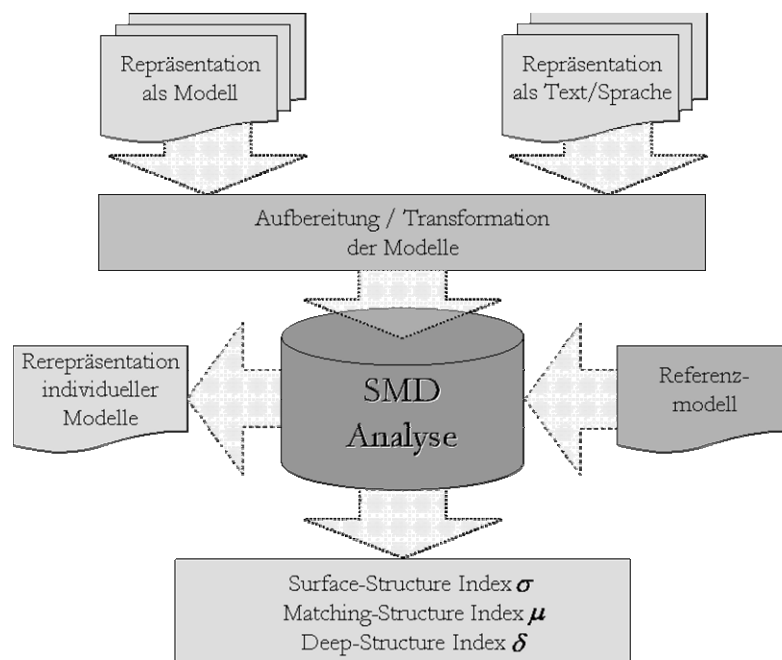


Abbildung 11: Die einzelnen Schritte der SMD-Technologie

Die Analyse der Deep-Structure erfordert zusätzlich ein Referenzmodell (vgl. Kapitel 4.4.3), welches auch in Listenform vorliegen soll. Jedem individuellen MODELL kann als Ergebnis der SMD Analyse ein Surface-, Matching-, und Deep-Structure Index zugewiesen werden. Nach Bedarf können die individuellen MODELLE mittels GraphViz (AT&T, 2006) rerepräsentiert werden (vgl. Abbildung 12).

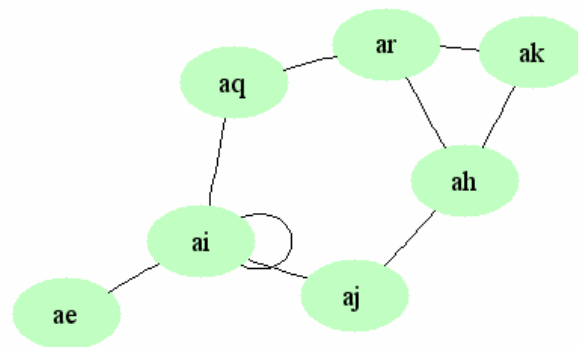


Abbildung 12: Rerepräsentation eines MODELLS der VPN 207 zum MZP 1

4.5.1 Manuelle Analyse

Eine manuelle Analyse durch einen Diagnostiker eignet sich aus zeitökonomischen Gründen nur bei einer geringen Datenmenge und wenig komplexen MODELLEN. Während der Entwicklung der *SMD-Technologie* wurde die manuelle Analyse mit mehreren Diagnostikern angewendet und auf Reliabilität überprüft. Die Ergebnisse in Tabelle 1 zeigen eine sehr hohe Zuverlässigkeit bzw. Übereinstimmung der Diagnostiker auf den Ebenen der Surface- und Matching-Structure. Der Übereinstimmungskoeffizient Kendall-W ($\tau = .749$) erwies sich auf der Ebene der Deep-Structure im Gegensatz zu den beiden anderen Ebenen als niedriger. Insgesamt zeigen die Ergebnisse jedoch eine hohe Reliabilität in Bezug auf die manuelle Analyse der *SMD-Technologie*. Diese hohe Reliabilität der manuellen Anwendung der *SMD-Technologie* erfordert jedoch eine umfangreiche Einweisung der Diagnostiker in die einzelnen Analyseschritte.

Tabelle 1: Reliabilität der manuellen Analyse der SMD-Technologie

	Surface- Structure	Matching- Structure	Deep- Structure
N Diagnostiker	2	2	2
Chi-Quadrat	253.914	272.029	209.726
df	140	140	140
Kendalls-W	.907	.972	.749
Asymptotische Signifikanz	.000	.000	.000

Die Analyseschritte gestalten sich im Einzelnen wie folgt:

1. Jedes MODELL wird durch den Diagnostiker einzeln begutachtet.
2. Diagnostiker müssen sich an die Richtsätze (Kapitel 4.3.1) strengstens halten.
3. Der Surface-Structure Index θ errechnet sich aus der Summe der Propositionen der einzelnen MODELLE (vgl. Kapitel 4.4.1).
4. Zur Berechnung des Matching-Structure Index μ werden die MODELLE in eine Abstandsmatrix übertragen. Anschließend kann μ mit der in Kapitel 4.4.2 dargestellten Formel berechnet werden.
5. Die Berechnung des Deep-Structure Index δ erfordert eine umfassende Analyse eines MODELLS bezüglich der semantischen Ähnlichkeit mit einem Referenzmodell. Die Ähnlichkeit zwischen einem MODELL und dem Referenzmodell wird als δ angegeben (vgl. Kapitel 4.4.3).

Der durchschnittliche Zeitaufwand für eine vollständige Analyse der Surface-, Matching- und Deep-Structure durch einen Diagnostiker beträgt bei MODELLEN mit max. 20 Propositionen und Referenzmodellen mit max. 20 Propositionen ca. 15 Minuten pro MODELL. Der hohe Zeitaufwand der manuellen Anwendung der *SMD-Technologie* führte schließlich zur Entwicklung und Implementation einer automatisierten Analyse.

4.5.2 Automatisierte Analyse

Mit steigender Anzahl der Propositionen eines MODELLS bzw. eines Referenzmodells und einer größeren Datenmenge ist eine manuelle Anwendung der *SMD-Technologie* aus zeitökonomischen Gründen nicht durchführbar. Die automatisierte Analyse der *SMD-Technologie* wurde webbasiert implementiert, wodurch eine plattformunabhängige Anwendung garantiert wird. Außerdem ist keine Installation einer zusätzlichen Software durch einen Diagnostiker notwendig. Der Zugriff auf die *SMD-Technologie* erfolgt durch simples Einwählen

mittels eines Internetbrowsers¹⁴. Die datenbankgestützte Anwendung der automatisierten *SMD-Technologie* wurde in PHP¹⁵ und Perl¹⁶ programmiert und greift auf eine MySQL-Datenbank¹⁷ zurück.

SMD TECHNOLOGY
RELATIONS | STRUCTURE | SEMANTICS

For analysing your data, please identify your Username, Referencemodel ID, Subject IDs, and Measurementpoint IDs.

Username: (1)lfenthaler

Referencemodel ID: (2)200a

First Subject ID: (3)201

Last Subject ID: (3)291

First Measurementpoint ID: (4)1

Last Measurementpoint ID: (4)7

START ANALYSIS (5)

Designed by Dirk lfenthaler

Abbildung 13: WEB-Interface der automatisierten SMD-Technologie

Die Berechnung der Surface-, Matching- und Deep-Structure Indizes wurde mittels PHP- und Perl-Code umgesetzt. Nach dem Import der in Listenform (vgl. Kapitel 4.3) vorliegenden MODELLE in die Datenbank kann sofort mit der Analyse begonnen werden (vgl. Abbildung 13):

¹⁴ Die automatisierte Analyse der SMD-Technologie ist mit den zurzeit gängigen Internetbrowsern (Microsoft Internet Explorer, Mozilla Firefox, etc.) kompatibel.

¹⁵ PHP ist eine Skriptsprache zur Erstellung von dynamischen Webseiten (www.php.net).

¹⁶ Perl ist eine plattformunabhängige Skriptsprache, welche als Synthese aus C und UNIX-Befehlen entstanden ist (www.perl.com).

¹⁷ MySQL (My Structured Query Language) ist eine frei verfügbare Datenbankanwendung zum Speichern von Information auf einem Server (www.mysql.com).

1. Der Diagnostiker kann einen freien Namen (Username) für die SMD Analyse vergeben, um später auf die Daten in der Datenbank zurückgreifen zu können.
2. Als nächstes wird die Nummer des Referenzmodells für die SMD Analyse eingegeben.
3. Sollen mehrere Versuchspersonen gleichzeitig analysiert werden, kann bei chronologischer Abfolge der VPN-Nummern die erste und letzte VPN-Nummer eingegeben werden. Die SMD Analyse wird dann für jede Versuchsperson einzeln durchgeführt.
4. Wurden pro Versuchsperson mehrere Messzeitpunkte (MZP) erhoben, können auch diese eingegeben und einzeln analysiert werden.
5. Mit drücken des Buttons „start_analysis“ wird der Rechengang der SMD Analyse gestartet.

Die Dauer der Analysephase ist von der Komplexität des Referenzmodells und der Versuchspersonenmodelle, sowie von der Anzahl der einzelnen Modelle abhängig. Eine vollständige Analyse aller im empirischen Teil dieser Arbeit ausgewerteten MODELLE (106 VPN á 7 Modelle) benötigt zirka 3 Minuten. Alle Ergebnisse der *SMD-Technologie* (Username, VPN-Nummer, MZP-Nummer, Surface-, Matching-, Deep-Structure, Auswertungszeit) werden in der SQL-Datenbank gespeichert und können in beliebiger Form ausgegeben und weiterverarbeitet werden. Zusätzlich wird eine dynamische HTML-Seite mit allen Ergebnissen inklusive einer Rerepräsentation der individuellen MODELLE generiert und automatisch angezeigt (vgl. Abbildung 14).

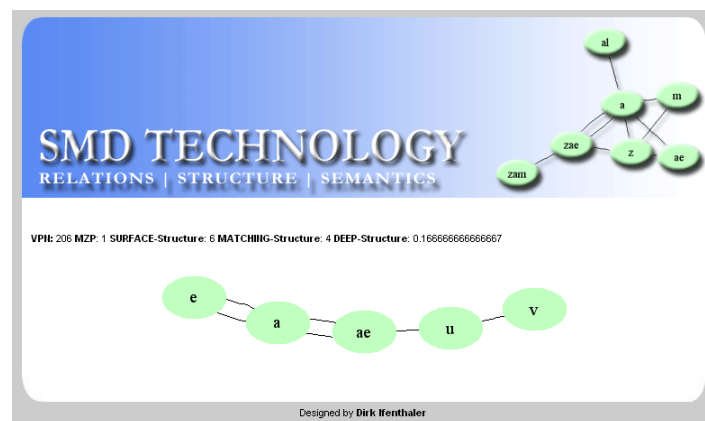


Abbildung 14: Ergebnisausgabe der automatisierten SMD-Technologie

4.6 ZUSAMMENFASSUNG DES DIAGNOSEVERFAHRENS

Das als *SMD-Technologie* vorgestellte Diagnoseverfahren ist in Anlehnung an die Theorie der mentalen Modelle (vgl. Johnson-Laird, 1983; Seel, 1991) entwickelt worden. Die umfangreichen Anforderungen an die *SMD-Technologie* ergeben sich aus der Tatsache, dass kognitive Konstrukte, wie das der mentalen Modelle, nicht unmittelbar diagnostizierbar sind.

Die *SMD-Technologie* wurde im Hinblick auf verschiedene Verfahren zur Erfassung von Wissensstrukturen (vgl. Kapitel 3) mit dem Ziel entwickelt, eine möglichst präzise, objektive, reliable und valide Diagnose externalisierter Modelle zu ermöglichen. Externalisierte Modelle können auf einer relationalen, strukturellen und semantisch-strukturellen Ebene begutachtet werden. Jede der drei Ebenen bildet ein von den anderen Ebenen unabhängiges quantitatives Maß ab, welches für weitere statistische Analysen verwendet werden kann. Durch die Domänenunabhängigkeit und Flexibilität der *SMD-Technologie* kann eine Diagnose in nahezu jedem Inhaltsbereich erfolgen. Daraus ergibt sich eine Vielzahl von Einsatzmöglichkeiten, sowohl im Forschungs- als auch im Anwendungsbereich.

Sowohl die manuelle als auch die automatisierte Anwendung der *SMD-Technologie* erwies sich bei der Analyse umfangreicher Datensätze (vgl. Kapitel 7f) als ein fehlerfreies Diagnoseinstrument. Dennoch hat sich die automatisierte Anwendung der *SMD-Technologie* durchgesetzt, da der Informationsgewinn identisch, die Reliabilität der Diagnose höher und der Zeitaufwand um ein vielfaches geringer einzuschätzen ist.

In Bezug auf die in dieser Arbeit gestellte Forschungsfrage nach einer Diagnose lernabhängiger Veränderung mentaler Modelle, ist neben der präzisen Erfassung mentaler Modelle auch die Frage einer testtheoretisch sauberen Prozessdiagnostik zu stellen. Das im folgenden Kapitel vorgestellte Verfahren der Veränderungsmessung soll die methodologische Grundlage für den empirischen Teil dieser Arbeit bilden.

5 VERÄNDERUNGSMESSUNG ALS VERFAHREN DER EMPIRISCHEN LEHR-LERN-FORSCHUNG

„Es ist nichts beständig als die Unbeständigkeit.“ (Immanuel Kant, dt. Philosoph, 1724-1804)

Die Diagnose von Veränderungen gehört nach wie vor zu den bedeutendsten Fragestellungen der empirisch orientierten Wissenschaften. Nach Eid (2003) ist die Aufgabe der Veränderungsmessung, Informationen über Erhalt und Veränderung von Verhalten zu analysieren und den methodologischen Rahmen hierfür zu stellen. Als besonders relevant stellt sich die Frage nach Veränderungen in den psychologischen Teildisziplinen heraus (vgl. Stieglitz, 1986). Veränderungsmessungen haben sowohl in der Entwicklungspsychologie (vgl. Rothgang, 2003) und Sozialpsychologie (vgl. Bierhoff, 2005) als auch in der anwendungsorientierten klinischen Psychologie (vgl. Perrez, 2005) einen zentralen Stellenwert. Darüber hinaus finden sich in der pädagogisch-psychologischen Forschung (vgl. Gage & Berliner, 1996) Fragestellungen, die eine Veränderungsmessung unumgänglich machen. Außerdem umfasst der Bereich Schul- und Unterrichtsforschung viele Fragestellungen mit Fokus auf Veränderungen (vgl. Doll & Prenzel, 2001, 2003). Ein weiterer Forschungsbereich mit Veränderungsfragestellungen stellt die universitäre Lehre (vgl. Doll & Knopf, 2001; Vögele, 2004; Wild, 2000) und die Erwachsenenbildung dar (vgl. Eckert & Tippelt, 2001).

Die Lehr-Lern-Forschung befasst sich insbesondere mit dem sukzessiven Aufbau und der Veränderung psychischer Dispositionen bei der Informationsverarbeitung, mit der didaktisch optimal aufbereiteten Vermittlung von Fertigkeiten und den komplexen Zusammenhängen zwischen Lehren und Lernen (vgl. Seel, 2003a). Die Theorie der mentalen Modelle (vgl. Johnson-Laird, 1983; Seel, 1991) wird diesem Forschungsbereich zugeordnet. Wie bereits in den vorangegangenen Kapiteln dieser Arbeit deutlich wurde, muss sich folglich die Diagnose lernabhängiger Veränderung mentaler Modelle aus methodologischer Sicht mit dem Verfahren der Veränderungsmessung auseinandersetzen.

Zunächst wird eine Einordnung und Begriffsbestimmung der in Bezug auf die Veränderungsmessung verwendeten Terminologie vorgenommen (Kapitel 5.1). An die Publikation von Harris (1963) reiht sich eine Vielzahl von Veröf-

fentlichungen an (vgl. Cronbach & Furby, 1970; Helmreich, 1977; Petermann, 1978; Krauth, 1981; Willett, 1989; Collins & Horn, 1991; Renkl & Gruber, 1995; Collins & Sayer, 2001; Moskowitz & Hershberger, 2002), die sich mit den methodologischen Fragen der Veränderungsmessung befassen. Diese Publikationen werden in einer historischen Synopse der Veränderungsmessung dargestellt (Kapitel 5.2). Die wissenschaftliche Diskussion über die Veränderungsmessung befasst sich auch immer wieder mit speziellen Problemen, auf die im Kapitel 5.3 eingegangen wird. Anschließend werden verschiedene Testmodelle vorgestellt (Kapitel 5.4). In Kapitel 5.5 über die Modellierung von Veränderung mit HLM (Raudenbusch & Bryk, 2002) wird schließlich die testtheoretische Grundlage für den empirischen Teil dieser Arbeit gelegt.

5.1 BEGRIFFSBESTIMMUNGEN

Die lernprozessbegleitende Diagnostik ist ein spezieller methodologischer Ansatz, der über besondere Terminologien verfügt. Die zentralen Termini *Veränderungsmessung*, *Variabilität* und *Veränderung*, *messbare Größe der Veränderung* und *interindividuelle Unterschiede in intraindividuellen Veränderungen* werden im Folgenden kurz bestimmt.

5.1.1 Veränderungsmessung

Um aus methodologischer Sicht eine Begriffsverwirrung bezüglich der *Veränderungsmessung* zu vermeiden, unterscheidet Pawlik (1976) zwischen einer Status- und einer Prozessdiagnostik. Die *Statusdiagnostik* erfasst den Status quo einer psychologischen Disposition, wobei davon ausgegangen wird, dass diese sowohl zeitlich als auch situativ unverändert bleiben. Demgegenüber beschreibt die *Prozessdiagnostik* zeitlich und situativ bedingte Veränderungen von psychologischen Dispositionen. Die Definition von Kleiter (1987, p. 235) soll eine Begriffsklarheit schaffen:

„Veränderungsmessung umfasst all diejenigen Datenerhebungs- und statistischen Rechenverfahren, bei denen mehrmalige Messungen, d.h. Meßwiederholungen in der Zeit zum gleichen Gegenstand erhoben werden.“

In experimentellen Bedingungen bleibt es dem Forscher überlassen, zwischen einzelnen Messzeitpunkten pädagogische Interventionen kontrolliert einzusetzen, um somit neue Erkenntnisse zu gewinnen. Allgemein lässt sich eine experimentelle Anordnung der Veränderungsmessung wie folgt darstellen (vgl. Abbildung 15).

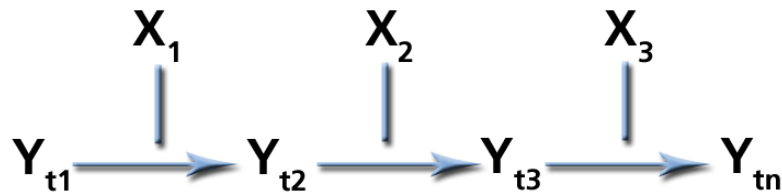


Abbildung 15: Allgemeine experimentelle Anordnung der Veränderungsmessung in Anlehnung an Kleiter (1987)

Die messbare Größe Y_j wird an einzelnen Messzeitpunkten t_n diagnostiziert. Abhängig von der experimentellen Anordnung können pädagogische Interventionen X_i eingesetzt werden. Die allgemeine experimentelle Anordnung lässt sich bezüglich der messbaren Größen, der Messzeitpunkte und der pädagogischen Intervention an den Untersuchungsgegenstand bzw. an die wissenschaftliche Fragestellung anpassen.

5.1.2 Variabilität und Veränderung

Im allgemeinen Sprachgebrauch werden für *Veränderung* bedeutungsähnliche Ausdrücke wie *Dynamik*, *Mobilität*, *Wandel* und *Variabilität* verwendet (vgl. Stieglitz, 1986; Renn, 1973). Im englischsprachigen Raum stehen die Vokabeln *change*, *growth* und *gain* für *Veränderung* (Helmreich, 1977). Als Gegenpol von *Veränderung* wird häufig der Begriff der *Konstanz* benutzt.

Die *Veränderungsmessung* als Verfahren der Prozessdiagnostik unterscheidet zwischen *Variabilität* und *Veränderung* (vgl. Eid, 2003). Die *Variabilität* wird als kurzfristiger und reversibler Wechsel psychologischer Dispositionen betrachtet.

Aus pädagogischer Sicht wäre eine *Variabilität* bezogen auf die kognitive Leistungsfähigkeit einer Person eine situationale Oszillation, die durch Stress o.ä. verursacht wurde. Da die *Variabilität* als reversibel betrachtet wird, kann davon ausgegangen werden, dass die kognitive Leistungsfähigkeit einer Person ohne den situativen Faktor Stress der eigentlichen psychologischen Disposition entspricht.

Veränderungen vollziehen sich gegenüber der *Variabilität* langsamer, sind situationsübergreifend und dauerhafter (vgl. Eid, 2003). *Veränderungen* werden als sehr schwer bzw. nicht reversibel betrachtet.

Wird durch eine pädagogische Intervention eine *Veränderung* der psychologischen Disposition einer Person hervorgerufen, so wird diese als situationsübergreifend betrachtet. Ein Grundschüler, der Lesen und Schreiben lernt, wird in den meisten Fällen dauerhaft diese Fähigkeiten besitzen und diese nicht wieder vergessen¹⁸.

Eid (2003) weist darauf hin, dass *Veränderungen* mit situationsbedingten *Variabilitäten* interferieren können. Daher erweist es sich bei einer Diagnose von *Veränderung* als notwendig, „die situationsbedingte Variabilität zu eliminieren“ (Eid, 2003, p. 108).

5.1.3 Messbare Größe der Veränderung

Helmreich (1977, p. 23) beschreibt zwei notwendige Bedingungen für die Veränderungsmessung:

1. Der Wert einer Variablen der Veränderungsmessung muss sich über mehrere Messzeitpunkte quantitativ verändern können.
2. Die Variable selbst soll dabei unverändert bleiben.

Wird eine Variable für die Diagnose von *Veränderung* eingesetzt, darf sich diese über mehrere Messzeitpunkte bezogen auf das zu messende Merkmal nicht verändern. Der Wert dieser Variablen soll sich jedoch quantitativ verändern können, da ohne diese Eigenschaft keine Veränderungsmessung erfolgen kann. Renn (1973, p. 50) beschreibt diese Forderung wie folgt:

“Nur wenn gewiß ist, daß sich die Bedeutung der Indikatoren im Zeitablauf nicht gewandelt hat, ist ein Vergleich der zu beiden Zeitpunkten gewonnen Meßwerte sinnvoll.“

Im Zusammenhang mit der messbaren Größe von *Veränderungen* muss bezüglich des eingesetzten Messinstrumentes beachtet werden, dass sich dieses

¹⁸ Es sind durchaus Fälle von funktionalem bzw. sekundärem Analphabetismus bekannt. Eigler (1990) berichtet, dass laut Schätzungen in Deutschland 2-3 Millionen funktionale Analphabeten leben, d.h. Personen, die Lese- und Schreibunterricht in der Schule erhielten und im Laufe der Zeit ihre Kompetenz in der Ausübung der Lese- und Schreibfähigkeit verloren.

im Zeitverlauf nicht verändert. Helmreich (1977, p. 24) weist auf Messinstrumente hin, die sich „verbrauchen“. Selbst einzelne Items können über mehrere Messzeitpunkte eine andere Bedeutung erhalten.

5.1.4 Interindividuelle Unterschiede in intraindividuellen Veränderungen

Werden eine oder mehrere psychologische Dispositionen einer Person zu verschiedenen Messzeitpunkten diagnostiziert, so bildet die *Veränderungsmessung* eine *intraindividuelle Veränderung* ab (vgl. Krauth, 1981). Viele Fragen der Lehr-Lern-Forschung gehen jedoch über diese einzelfallanalytische Sichtweise hinaus. Häufiges Forschungsinteresse ist außerdem die Diagnose *interindividueller Unterschiede* in *intraindividuellen Veränderungen* (vgl. Eid, 2003, p. 111). Die unterschiedlichen Fragestellungen der *Veränderungsmessung* erfordern differenzierte experimentelle Bedingungen und müssen bezüglich der methodologischen Auswertungsstrategie mit berücksichtigt werden.

5.2 HISTORISCHE SYNOPSE DER VERÄNDERUNGSMESSUNG

Fahrenberg (1968, pp. 42) berichtet über die Anfänge der psychologischen Verlaufsanalyse. Oehrn (1889) und Seashore & Kent (1905) seien die ersten, die von einer quantitativ beobachteten intraindividuellen Variabilität berichten. Die frühen empirischen Arbeiten zeigen jedoch, bedingt durch „[...] die geringe Zahl und einseitige Auswahl der untersuchten Variablen [...]“ und durch „[...] die mathematisch-statistischen Schwierigkeiten [...]“, ihre Grenzen auf (Fahrenberg, 1968, p. 44). Die Verlaufs-forschung wurde durch Publikationen von Persönlichkeitsforschern um Heiß und Cattell fortgeführt¹⁹.

Die Veröffentlichung des von Harris (1963) herausgegebenen Sammelbandes „Problems in Measuring Change“ misst der methodologischen Diskus-

¹⁹ Nach Fahrenberg (1968, p. 44) veröffentlichten Heiß und Cattell im Jahre 1947 gleichzeitig Beiträge zur Persönlichkeitspsychologie, die sich mit der Verlaufs-forschung befassten.

sion um die *Veränderungsmessung* eine große Bedeutung bei²⁰. Bereiter (1963) befasst sich in seinem Grundlagenartikel mit Fragen zum „over-correction-under-correction dilemma“, „unreliability-invalidity dilemma“ und „physicalism-subjectivism dilemma“. Neben univariaten und multivariaten Analysemodellen (vgl. Gaito & Wiley, 1963; Bock, 1963; Horst, 1963) werden in Harris (1963) auch statistische Modelle für Einzelfallanalysen diskutiert (vgl. Holtzmann, 1963).

Cronbach & Furby (1970, p. 68) greifen die Diskussion um die *Veränderungsmessung* auf und bezeichnen diese als „a persistent puzzle in psychometrics“. Die Methodenentwicklung für die Erfassung von Veränderungen bildet ein weiterer Schwerpunkt der Publikationen. Fahrenberg et al. (1977) beleuchten neben einer Vielzahl unterschiedlichster Methoden auch die testmethodischen Probleme von wiederholten Beobachtungen. Die Veränderungsmessung im Rahmen der klassischen Testtheorie wird von Petermann et al. (1977) diskutiert. Das umfassende Methodenspektrum von Petermann (1978) führt beispielhaft in *probabilistische Testmodelle*, *Makroff-Modelle*, *Computersimulationen* und *Einzelfallbetrachtungen* ein. Tack (1980) nimmt eine Formalisierung der Veränderungsmessung vor, um „[...] insbesondere Probleme der Situationsabhängigkeit von Meßergebnissen im Rahmen der Veränderungsmessung klarer zu sehen [...]“ (Tack, 1980, p. 105). Fischer & Formann (1982) besprechen ein linear-logistisches Modell der Veränderungsmessung. Das LLRA²¹ erfordert keine Quantifizierung der Beobachtungen, liefert dennoch quantitative Effektparameter. Für Spada (1983, p. 83) stellt sich die Frage, inwieweit „die Erfassung von Veränderungen ein trauriges Kapitel“ ist.

Ausgehend von einer Literaturanalyse gibt Petermann (1986) einen Überblick über Probleme und neuere Entwicklungen der Veränderungsmessung. Die Schwerpunkte dieser Arbeit liegen in den konzeptuellen Schwierigkeiten der Veränderungsmessung, der Testtheorie, der Einzelfalldiagnostik und Einzelfallstatistik und formalen Modellen wie LISREL und Computersimulationen.

²⁰ Der Sammelband ist das Ergebnis einer dreitägigen Konferenz, die 1962 in Madison, Wisconsin durchgeführt wurde. Ziel der Konferenz war es, die aktuellen Probleme der Veränderungsmessung zu diskutieren und kritisch zu hinterfragen (vgl. Harris, 1963, p. vii).

²¹ LLRA steht für „linear logistic model with relaxed assumptions“ (vgl. Fischer & Formann, 1982).

nen. Kleiter (1987) diskutiert Stand und Perspektiven der Veränderungsmessung aus formal-gegenständlicher Sicht. Willett (1988, 1989) beklagt, dass die Vielzahl an Publikationen über Veränderungsmessungen nur wenig zur Lösung empirischer Fragestellungen bezüglich des Lernens beigetragen haben.

Collins & Horn (1991) stellen „best methods for the analysis of change“ vor. Darin werden aktuelle Fortschritte, unbeantwortete Fragen und zukünftige Richtungen der Veränderungsmessung vorgestellt. Die Möglichkeiten einer adäquaten statistischen Modellierung von Veränderungen mittels Hierarchisch Linearer Modelle verdeutlichen Renkl & Gruber (1995) am Beispiel der Gedächtnispsychologie. Collins & Sayer (2001) berichten, dass seit der Publikation von Collins & Horn (1991) enorme Fortschritte im Bereich der Veränderungsmessung gemacht wurden. Mit dem Sammelband von Moskowitz & Hershberger (2002) wird der aktuellen Diskussion um Methoden und Anwendungen der Veränderungsmessung Nachdruck verliehen. Die AEPF²² legte in ihrer Frühjahrstagung 2005 ihren Fokus auf die Veränderungsmessung und Längsschnittstudien. Die Tagung an der Freien Universität Berlin fand mit 23 Symposien, 83 Einzelvorträgen und 25 Postern eine „ungewöhnliche Resonanz“ (AEPF, 2005, p. 7). Neben aktuellen Forschungsstudien wurden spezielle Probleme der Veränderungsmessung diskutiert, die im Folgenden näher beleuchtet werden.

5.3 SPEZIELLE PROBLEME DER VERÄNDERUNGSMESSUNG

Ausgehend von der klassischen Testtheorie wird angenommen, dass ein kognitives Konstrukt durch eine gemessene Größe x_i beschrieben werden kann. Diese setzt sich aus der wahren Größe X_i und dem dazugehörigen Messfehler e_{xi} zusammen. Demzufolge wird die gemessene Größe additiv aus dem wahren Wert und dem dazugehörigen Fehler der Messung bestimmt (vgl. Petermann; 1978, p. 26). Aus mathematischer Sicht liegt es auf der Hand, Veränderungen eines kognitiven Konstruktes zwischen zwei Messzeitpunkten durch

²² AEPF ist die Arbeitsgruppe empirisch-pädagogische Forschung.

ein Differenzmaß abzubilden (vgl. Stelzl, 2005). Neben einfachen Differenzmaßen werden Korrelations-, Regressions- und Residualmaße unterschieden (vgl. Petermann, 1978, pp. 33). Folgt man den Argumentationen von Bereiter (1963, p. 3), dürfen jedoch drei „persisting dilemmas in the measurement of change“ nicht außer Acht gelassen werden. Diese sind:

1. The over-correction-under-correction dilemma (Regressionseffekt)
2. The unreliability-invalidity dilemma (Reliabilitäts-Validitäts-Dilemma)
3. The physicalism-subjectivism dilemma (Messbedeutungsproblem)

5.3.1 Regressionseffekt

Ein Regressionseffekt, auch Regression zur Mitte genannt, kann dann auftreten, wenn Veränderungsmessungen mit identischen Messinstrumenten durchgeführt werden. Bedingt durch die statistische Regression²³ neigen Extremwerte bei wiederholten Messungen dazu, sich dem Mittelwert der Verteilung anzunähern. Dies bedeutet, dass hohe Werte der ersten Messung niedrigere Werte in der folgenden Messung erzielen (losers) und niedrige Werte der ersten Messung höhere Werte erreichen (gainers). Folglich kann der Regressionseffekt speziell bei der Diagnose von Extremgruppen zu einer verfälschten Interpretation der Ergebnisse führen.

Die vereinfachte Darstellung in Abbildung 16 illustriert den Regressionseffekt am Beispiel eines Wissenstest.

Bedingt durch die Messfehlereinflüsse ist es wahrscheinlich, dass ein Wert des ersten Messzeitpunktes (MZP1) um eine bestimmte Größe über- bzw. unterbewertet wird. Wird der Wissenstest an einem zweiten Messzeitpunkt (MZP2) wiederholt, so ist es unwahrscheinlich, dass die Messfehlereinflüsse unverändert bleiben. Folglich tendieren Probanden, die beim ersten Messzeitpunkt einen hohen Wert im Wissenstest erreicht haben, beim zweiten Messzeitpunkt zu einem niedrigeren Wert. Probanden mit einem niedrigen Wert beim ersten Messzeitpunkt erzielen infolgedessen einen höheren Wert beim zweiten Messzeitpunkt. In beiden Fällen ist es wahrscheinlich, dass Messfehler den hohen bzw. niedrigen Wert beim ersten Messzeitpunkt herbeigeführt haben. Folgen weitere

²³ Die statistische Regression bezeichnet die auf Messfehler zurückzuführende Tendenz zur Mitte (vgl. Petermann, 1978, p. 28).

Messzeitpunkte, werden die Werte entweder ihren wahren Werten näher kommen oder in das andere Extrem tendieren.

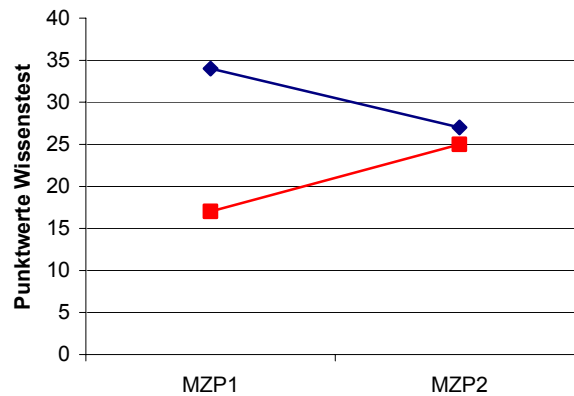


Abbildung 16: Der Regressionseffekt am Beispiel eines Wissenstest (vgl. Petermann, 1978, p. 29)

Werden weitere Variablen berücksichtigt, die mit der veränderten Variable in Zusammenhang stehen, werden die Ergebnisse durch den Regressionseffekt stärker verzerrt (vgl. Petermann et al., 1977, p. 135). Klauer (2005) weist darauf hin, dass der Regressionseffekt mit anderen Effekten, wie z. B. dem *Ceiling-Effekt*²⁴, kumulieren kann und daher weitaus schwerer zu kontrollieren wird.

Ein Regressionseffekt kann trotz eines sehr reliablen Tests nie ausgeschlossen werden. Aus diesem Grund werden in experimentellen Designs Kontrollgruppen herangezogen, die ausschließlich die verwendeten Tests bearbeiten. Bei dieser Kontrollgruppenanordnung ist zwar auch ein Regressionseffekt zu erwarten, jedoch kann davon ausgegangen werden, dass der Regressionseffekt sowohl in der Experimental- als auch in der Kontrollgruppe konstant gehalten und somit kontrolliert wird (vgl. Campbell & Stanley, 1966).

In experimentellen Anordnungen der Veränderungsmessung muss der Regressionseffekt folglich bereits bei der Planung eines Experimentes berücksichtigt werden. Aber auch bei der Interpretation der Ergebnisse ist der Regressionseffekt nicht außer Acht zu lassen, da hier die Gefahr besteht, dass

²⁴ Der Ceiling- oder Deckeneffekt beschreibt das Phänomen, dass hohe Messwerte in einer begrenzten Skala wenig dazu gewinnen können, da sich diese bereits am Maximum der jeweiligen Skala befinden (vgl. Klauer, 2005, p. 89).

selbst ein geringer Regressionseffekt mögliche Lerneffekte einer pädagogischen Intervention verfälscht.

5.3.2 Reliabilitäts-Validitäts-Dilemma

Ausgehend von der Annahme, dass durch eine Korrelation zwischen zwei Variablen geprüft werden kann, ob diese identische Konstrukte messen, schließt sich das Reliabilitäts-Validitäts-Dilemma der Veränderungsmessung an (vgl. Rost, 2004, pp. 280). Eine niedrige Korrelation bei einer Veränderungsmessung über mehrere Messzeitpunkte kann zum einen eine Veränderung des zu messenden Konstruktes bedeuten (z. B. ein Lernzuwachs). Andererseits kann die niedrige Korrelation aber auch ein verändertes Messinstrument bedeuten, d.h. es wird zu den unterschiedlichen Messzeitpunkten etwas Unterschiedliches gemessen (vgl. Al-Diban, 2002, p. 57).

Die Reliabilität von Differenzwerten wird von Lord (1963) wie folgt dargestellt:

$$r_{gg'} = \frac{s_y^2 r_{yy'} - 2s_y s_x r_{xy} + s_x^2 r_{xx'}}{s_y^2 - 2s_y s_x r_{xy} + s_x^2} \quad [5.3.1]$$

Geht man von einer einheitlichen Varianz ($s_x = s_y$) durch z-transformierte Variablen aus, so wird die Formel wie folgt vereinfacht (vgl. Lord, 1963; Petermann, 1978):

$$r_{gg'} = \frac{r_{yy'} + r_{xx'} - 2r_{xy}}{2(1 - r_{xy})} \quad [5.3.2]$$

Dabei ist $r_{gg'}$ die Reliabilität der Differenzen, $r_{xx'}$ die Reliabilität der ersten Messung, $r_{yy'}$ die Reliabilität der zweiten Messung und r_{xy} die Korrelation zwischen erster und zweiter Messung. Das nachstehende Praxisbeispiel soll die Reliabilitäts-Validitäts-Problematik verdeutlichen:

Ein Wissenstest, der an zwei Messzeitpunkten durchgeführt wurde, weist eine Reliabilität der ersten Messung von $r_{xx'} = .86$ und eine Reliabilität der zweiten Messung von $r_{yy'} = .82$ auf. Zwischen der ersten und zweiten Messung wurde eine niedrige Korrelation von $r_{xy} = .25$ berechnet. Aus der oben genannten Formel lässt sich eine Reliabilität der Differen-

zen von $r_{gg'} = .78$ berechnen. Die hohe Reliabilität der Differenzen ergibt sich jedoch nur, wenn die Korrelation zwischen der ersten und zweiten Messung niedrig ist. Das Ergebnis kann (1) als eine Veränderung des Wissens vom ersten zum zweiten Messzeitpunkt, (2) als eine Veränderung des Wissenstests oder (3) als eine Veränderung des Wissens und des Wissenstests interpretiert werden. Würde eine hohe Korrelation zwischen den beiden Messungen vorliegen, wirkt sich dies negativ auf die Reliabilität der Differenzen aus.

Das Reliabilitäts-Validitäts-Dilemma kann aus Sicht der klassischen Testtheorie nicht gelöst werden, wenn ein Messinstrument über mehrere Zeitpunkte identische Konstrukte messen soll und zusätzlich die Korrelation i.S. der Retest-Reliabilität mit in die Berechnung einfließt (vgl. Petermann et al., 1977). Helmreich (1977, p. 47) folgert, „[...] die Korrelation der Meßreihen von erster und zweiter Testung ist in diesem Fall nicht das adäquate Mittel, etwas über die Qualität von Tests auszusagen, die zur Messung der Zustände zum ersten und zweiten Meßzeitpunkt eingesetzt werden.“ Zusammenfassend kann festgestellt werden, dass bei einem Test die Validität sinkt, wenn die Retest-Reliabilität steigt (vgl. Rost, 2004, p. 393).

5.3.3 Messbedeutungsproblem

„A choice must be made between objective measures of change that are meaningless within the province of differential psychology and meaningful measures that carry the stigma of subjectivity.“ (Bereiter, 1963, p. 20)

Das dritte Dilemma der Veränderungsmessung wird von Bereiter (1963) als Physikalismus-Subjektivismus-Dilemma bzw. als Messbedeutungsproblem (vgl. Petermann, 1978) bezeichnet. Aus Sicht der beiden Autoren ist es nahezu unmöglich, präzise quantifizierbare Merkmale zu definieren, die auch aus pädagogisch-psychologischer Sicht inhaltlich sinnvoll interpretiert werden können.

In einer Unterrichtseinheit sollen Schüler bezüglich deren Lesefehler und -geschwindigkeit diagnostiziert werden. Vor und nach einer speziellen pädagogischen Intervention, die verschiedene Lesetechniken aufzeigt, werden die Lesefehler und -geschwindigkeiten der Schüler gemessen. Ein durchaus denkbare Ergebnis wäre, dass sich zwei Schüler im selben Umfang verbessern. Objektiv weisen beide Schüler die gleiche Veränderung auf. Aus pädagogisch-psychologischer Sicht kann diese Veränderung jedoch eine vollkommen andere Ursache haben. Daher ist es not-

wendig, die objektiv nicht erfassbaren Ursachen zu diagnostizieren, um die Lesekompetenz der Schüler individuell zu fördern.

Neben der objektiven Erfassung von Veränderung schlägt Bereiter (1963) eine subjektive Erfassung von Veränderung mit speziellen Skalen vor. Diese subjektiven Einschätzungsskalen sind jedoch für Experimente mit kurzen Zeitabständen unbrauchbar. Hinzuzufügen sind weitere Nachteile einer subjektiven Erfassung von Veränderung, wie z. B. die unklare Reliabilität, Skalierungsprobleme, Veränderung des Bezugssystems und die Abhängigkeit der Gedächtnisleistung von Probanden. Entscheidender Nachteil der subjektiven Erfassung von Veränderung ist nach Petermann (1978, p. 33), „[...] daß die Bedeutung der Wirkgrößen und der Veränderungsbetrag das Ergebnis von subjektiven Beurteilungen und dadurch »willkürlich« ist.“

5.3.4 Dilemmata einer Veränderungsmessung mentaler Modelle

Im Hinblick auf die Theorie der mentalen Modelle ergeben sich für eine Veränderungsmessung mentaler Modelle weitere Dilemmata (vgl. Al-Diban, 2002). Die empirische Lehr-Lern-Forschung nutzt zur Erfassung von Veränderung eine Vielzahl an Instrumenten in Form von mündlichen und schriftlichen Befragungen, systematischen Beobachtungen und physiologischen Messungen (vgl. Bortz & Döring, 1995; Kleber, 1992). Dabei ist der Forschungsgegenstand und die Fragestellung schließlich entscheidend, welches dieser unterschiedlichsten Instrumente Anwendung findet.

Der Forschungsgegenstand der mentalen Modelle und die Fragestellung nach einer lernabhängigen Veränderung mentaler Modelle schränken die zur Verfügung stehenden Instrumente ein. Die Tatsache, dass mentale Modelle nicht direkt diagnostiziert werden können, schließt Instrumente der Beobachtung ohne Zweifel aus. Es gilt weiter hinzuzufügen, dass mentale Modelle rein situativ in komplexen Problemlösesituationen konstruiert werden, was eine Verwendung von Fragebogeninstrumenten ausschließt.

Es ist festzuhalten, dass neben den unter 5.3 genannten speziellen Problemen der Veränderungsmessung, weitere Probleme bei der Diagnose mentaler Modelle zu berücksichtigen sind. So kann (1) eine Diagnose mentaler Modelle ausschließlich mittels geeigneter abhängiger Variablen erfolgen. Hierfür ist (2)

eine experimentelle Anordnung durch den Wissenschaftler einzurichten, die (3) die Probanden in eine komplexe, problemhaltige Situation versetzt, welche die Konstruktion, Revision und Reorganisation von mentalen Modellen ermöglicht (vgl. Al-Diban, 2002) und (4) über mehrere Messzeitpunkte erfasst.

5.4 TESTTHEORETISCHE MODELLE DER VERÄNDERUNGSMESSUNG

Ausgehend von den unter 5.3 diskutierten Problemen der Veränderungsmessung stehen neuere Methoden für die Analyse von Veränderungsdaten zur Verfügung. Nachstehende Forderungen an eine moderne Veränderungsmessung werden ins Besondere von Fischer & Formann (1982, p. 80) diskutiert:

- *Decken- und Bodeneffekte* sollen ausgeschlossen werden
- *Regressionseffekte* sollen ausgeschaltet oder jedenfalls kontrolliert werden
- Willkürliche *Verteilungsannahmen*, die zu Methodenartefakten führen können, sollen vermieden werden

Vor dem Hintergrund der klassischen Probleme der Veränderungsmessung und den Forderungen an eine moderne Veränderungsmessung wurde eine Vielzahl testtheoretischer Modelle entwickelt. Im Sammelband über statistische Methoden für die Längsschnittforschung berichtet von Eye (1990a, 1990b), dass jedoch eine große Lücke zwischen bereits entwickelten testtheoretischen Modellen und deren adäquaten Anwendung in unterschiedlichen Forschungskontexten bestehe. Gründe hierfür sind u.a. der Mangel an leistungstarker Computersoftware, die unzureichende Dokumentation der Computersoftware und die nicht ausreichenden mathematisch-statistischen Kenntnisse über vorhandene testtheoretische Modelle. Dennoch haben sich verschiedene testtheoretische Modelle als aussichtsreiche Analysemethoden für Veränderungsdaten herausgestellt.

Lineare Strukturgleichungsmodelle oder *Kausalmodelle* basieren auf Überlegungen der Pfad-, Regressions- und Faktorenanalyse (vgl. Bortz, 1993; Backhaus et al., 2000). Vor der Analyse werden Kausalhypothesen formuliert, die mittels Computerprogrammen wie LISREL (Jöreskog & Sörbom, 1989; Jöreskog et al., 1999), LISCOMP (Muthen, 1987) bzw. EQS (Bentler, 1986, 1992) geprüft

werden. Die Anforderungen an das Datenniveau (Intervallskala) und die notwendige Stichprobengröße ($N > 200$) stellen für viele empirische Untersuchungen jedoch eine unüberwindbare Hürde dar. Es gilt hinzuzufügen, dass Veränderungen mittels Strukturgleichungsmodellen lediglich auf Stichprobenebene und nicht auf Personenebene analysiert werden (vgl. Langer, 2004, p. 223).

Al-Diban (2002, p. 60) berichtet von drei testtheoretischen Modellen, die „für eine angemessene Darstellung von Veränderungen“ mentaler Modelle Anwendung finden können. Auch Rost (2004) und Tent & Stelzl (1993) diskutieren diese testtheoretischen Modelle.

Das *dreifaktorielle Rasch-Modell* ist ein probabilistisches Verfahren, welches von einer dreidimensionalen Datenstruktur ausgeht. Diese ergibt sich aus Personen, Items und Messzeitpunkten bzw. Situationen. Gleichung 5.4.1 zeigt ein Modell, welches Wechselwirkungen zwischen Personen und Items zulässt.

$$p(X_{vit} = 1) = \frac{\exp(\alpha_{vit})}{1 + \exp(\alpha_{vit})} \quad [5.4.1]$$

Dabei gilt für α_{vit} die additive Zerlegung aus Gleichung 5.4.2:

$$\alpha_{vit} = \sigma_i + \delta_{vt} \quad [5.4.2]$$

Die Ausprägung der Personeneigenschaft v zum Zeitpunkt t wird durch doppelt indizierte Parameter δ_{vt} beschrieben. Die Anwendung des *dreifaktoriellen Rasch-Modells* ermöglicht somit die Beschreibung individueller Veränderungen über mehrere Messzeitpunkte. Da es sich im beschriebenen Modell um eine additive Zerlegung handelt, sollte beachtet werden, dass die personenspezifischen Veränderungen nur interpretierbar sind, wenn der Itemparameter σ_i konstant gehalten wird.

Fischer & Formann (1982) berichten über eine Veränderungsmessung mittels *linear-logistischer Modelle* welche als Spezialfall des allgemeinen Rasch-Modells verortet werden können. Das Ziel des testtheoretischen Ansatzes ist es, Effekte unterschiedlicher Maßnahmen zwischen verschiedenen Gruppen von Personen zu vergleichen (vgl. Rost, 2004). Gleichung 5.4.3 zeigt das linear-logistische Testmodell für den Zeitpunkt t_1 :

$$P_{vi1} = P(+ | v, i, t_1) = \frac{\exp(\theta_{vi})}{1 + \exp(\theta_{vi})} \quad [5.4.3]$$

Dabei gibt Parameter θ_{vi} die Ausprägung eines Merkmals der jeweiligen Person an. Das Veränderungsmaß dieses testtheoretischen Ansatzes wird *global abgebildet*, was bedeutet, dass die getestete Maßnahme, z. B. ein Lernzuwachs, für alle Personen und Items in gleichem Maße beschrieben wird (vgl. Rost, 2004; Tent & Stelzl, 1993). Al-Diban (2002, p. 60) berichtet von Nachteilen, die mit der Anwendung des *linear-logistischen-Modells* verbunden sind. So erweist sich die Durchführung dieses testtheoretischen Modells als aufwändig. Zusätzliche Probleme ergeben sich durch häufig unvollständige Datenmatrizen.

Im weiteren Sinne können Veränderungshypothesen auch mit *Latent-Class-Modellen* geprüft werden (vgl. Tent & Stelzl, 1993). Veränderungen i.S. von Lernen werden dabei als Übergänge von einer latenten Klasse in eine andere angenommen. Dieses probabilistische Verfahren geht davon aus, dass festgelegte Antwortmuster in einem Fragebogen oder Test erwartet werden können (vgl. Rost, 1996). Eine mathematische Auseinandersetzung mit *Latent-Class-Modellen* führen Langeheine & Rost (1996) bzw. van de Pol et al. (1991).

Für eine weitere Vertiefung in testtheoretische Modelle für Veränderungsdaten sei an dieser Stelle auf Arbeiten von Petermann (1986), Engel & Reincke (1996), von Eye (1990a, 1990b), Goldstein (1979) Bartholomew (1967) und Ditton (1998) verwiesen²⁵. Neben den bereits vorgestellten testtheoretischen Modellen soll im Folgenden Kapitel die *Hierarchisch Lineare Modellierung* als ein weiterer Ansatz zur Modellierung von Veränderungsdaten vorgestellt werden.

5.5 MODELLIERUNG VON VERÄNDERUNG MIT HLM

Ausgehend von den unter 5.3 genannten Dilemmata der Veränderungsmessung mentaler Modelle und weiteren unter 5.4 diskutierten Modellen der

²⁵ Die genannten Arbeiten fassen aus Sicht des Autors wichtige Aspekte testtheoretischer Modelle zusammen. In diesen Arbeiten werden weitere Quellen testtheoretischer Modelle genannt, die der interessierte Leser konsultieren möge.

Veränderungsmessung stellt die *Hierarchisch Lineare Modellierung* (HLM) von Bryk und Raudenbush (1987, 1992) einen weiteren aussichtsreichen Ansatz zur Modellierung individueller Veränderungen dar (siehe auch Bryk, Raudenbush & Congdon, 1994, Raudenbush & Bryk, 2002, Raudenbush et al., 2004).

Die Diagnose lernabhängiger Veränderung mentaler Modelle erfordert es, mehrere Messzeitpunkte zu berücksichtigen, was durch eine Modellierung individueller Wachstumskurven mittels HLM möglich wird. Das mathematisch-statistische Verfahren von HLM kann als Erweiterung der multiplen Regression verstanden werden, wobei die Grundlage des Verfahrens ein Mehrebenen-Regressionsmodell darstellt. Wie der Name *Hierarchisch Lineare Modellierung* verdeutlicht, wird bei diesem Verfahren von einer hierarchischen Datenstruktur ausgegangen. Die hierarchischen Ebenen für Veränderungsmessungen werden in HLM wie folgt gebildet. Bei wiederholten Messungen ist der Prädiktor der ersten Ebene (Level 1) die Zeit bzw. der jeweilige Messzeitpunkt. Die zweite Ebene (Level 2) untersucht die Ebene der Individuen, die dritte Ebene (Level 3) die der Aggregateinheiten (experimentelle Gruppen). Die folgenden Gleichungen sollen die für eine Veränderungsmessung mittels HLM allgemein getroffenen Annahmen veranschaulichen (vgl. Raudenbush & Bryk, 2002, p. 162).

$$Y_{ti} = \pi_{0i} + \pi_{1i}\alpha_{ti} + \pi_{2i}\alpha_{ti}^2 + \dots + \pi_{pi}\alpha_{ti}^p + e_{ti} \quad [5.5.1]$$

Für Gleichung 5.5.1 gilt:

Y_{ti} das Merkmal der Versuchsperson i zum Messzeitpunkt t

π_{0i} das Ausgangsniveau der Versuchsperson i zum Messzeitpunkt 0

π_{pi} der Veränderungsparameter p für Versuchsperson i

α_{ti} die Zeitvariable für Versuchsperson i zum Messzeitpunkt t

e_{ti} das Residuum der Versuchsperson i zum Messzeitpunkt t

Ein besonders Merkmal von HLM ist die Annahme, dass die Veränderungsparameter interindividuell variieren. Somit wird in HLM auf der zweiten Ebene (Level 2) getestet, ob sich Individuen hinsichtlich deren Ausgangsparametern π_0 , deren linearen Veränderungsparametern π_1 und deren quadratischen Veränderungsparametern π_2 unterscheiden (vgl. Hox, 2002; Singer &

Willett, 2003). Die Unterschiede werden auf der zweiten Ebene auf experimentelle Variationen oder auf Personenmerkmale zurückgeführt (vgl. Raudenbush & Bryk, 2002; Renkl & Gruber, 1995).

$$\pi_{pi} = \beta_{p0} + \sum_{q=1}^{Q_p} \beta_{pq} X_{qi} + r_{pi} \quad [5.5.2]$$

Für Gleichung 5.5.2 gilt:

X_{qi} das Versuchspersonenmerkmal (z. B. Geschlecht) bzw. die *dummy-Variable*, welche eine experimentelle Variation (z. B. Lernhilfen) repräsentiert

β_{pq} der Effekt von X_q auf den q -ten Veränderungsparameter

r_{pi} das Residuum für Versuchsperson i

Für die Diagnose lernabhängiger Veränderung mentaler Modelle eignet sich der Einsatz von HLM aus mehreren Gründen (vgl. Hofmann et al., 1993). Das mehrbenenanalytische Verfahren berücksichtigt (1) individuell verschiedene Veränderungsprozesse. Daher können nicht nur gruppenspezifische Charakteristika identifiziert werden, sondern auch individuelle Verlaufskurven diagnostiziert werden (vgl. Renkl & Gruber, 1995). Die unter 5.3 diskutierten Dilemmata durch (2) zeitliche Abhängigkeiten werden mit HLM adäquat modelliert (vgl. Eid, 2003). Durch hinzufügen einer weiteren Ebene (Level 3) kann neben Messzeitpunkten und Versuchspersonen (3) eine Ebene höherer Ordnung (z. B. Experimentalgruppe, Schule, etc.) berücksichtigt werden (vgl. Langer, 2004; Singer & Willett, 2003). Im Hinblick auf die zentrale Fragestellung nach einer lernabhängigen Diagnose mentaler Modelle, sollen durch die Analyse der vorliegenden Daten mittels HLM „vor allem grundlegende theoretische Fragen geklärt werden können“ (Renkl & Gruber 1995, p. 189).

6 FRAGESTELLUNGEN UND HYPOTHESEN

Ausgehend von der Theorie der mentalen Modelle (vgl. Seel, 1991) wurden im theoretischen Teil dieser Arbeit verschiedene Verfahren zur Wissensdiagnose hinsichtlich deren Eignung für eine Längsschnittuntersuchung beleuchtet. Darüber hinaus wurde die *SMD-Technologie* als neues Verfahren für die Diagnose lernabhängiger Veränderungen externalisierter Modelle entwickelt. Aus methodologischer Sicht konnten die Probleme einer Veränderungsmessung dargestellt werden, die bei Untersuchungen mit Längsschnittdesign berücksichtigt werden müssen. Diese theoretischen Annahmen (Kapitel 2 bis 5) sollen im folgenden, empirischen Teil dieser Arbeit überprüft und kritisch diskutiert werden.

Die zentrale Fragestellung nach einer *präzisen Diagnose lernabhängiger Veränderung mentaler Modelle* (vgl. Seel, 1997, 1999a) kann jedoch nur dann überprüft werden, wenn die methodologischen Rahmenbedingungen exakt auf die theoretischen Annahmen abgestimmt werden. Folglich ist ein Untersuchungsdesign zu entwickeln, welches die Konstruktion mentaler Modelle erfordert, die Externalisierung mentaler Modelle durch Probanden ermöglicht und eine längsschnittliche Datenanalyse integriert. Darüber hinaus werden reliable und valide Instrumente für eine umfassende und präzise Diagnose lernabhängiger Veränderung mentaler Modelle benötigt. Aus pädagogischer Sicht schließt sich die Frage an, ob spezielle Interventionen die Konstruktion mentaler Modelle beeinflussen.

Die einzelnen Fragestellungen und Hypothesen lassen sich den Bereichen Diagnose lernabhängiger Veränderungen (Kapitel 6.1), pädagogische Interventionen (Kapitel 6.2) und Effekte von Lernervoraussetzungen auf die Konstruktion mentaler Modelle (Kapitel 6.3) zuordnen.

6.1 DIAGNOSE LERNABHÄNGIGER VERÄNDERUNGEN

Nach Johnson-Laird (1989) vollzieht sich die lernabhängige Veränderung mentaler Modelle von *Novizenmodellen* hin zu *Expertenmodellen*. Anknüpfend daran sieht Snow (1990) die Veränderung mentaler Modelle zwischen *Präkon-*

zeptionen oder *Misskonzeptionen*, die als anfängliche mentale Modelle definiert werden, hin zu *kausalen Erklärungen*, die als angestrebte Endzustände des Lernens betrachtet werden (vgl. Seel et al., 2000). Folglich werden erhebliche Unterschiede zwischen Experten- und Novizenmodellen, aber auch zwischen individuellen Modellen während eines Lernprozesses erwartet. Aus diesen Annahmen lässt sich die erste übergeordnete Fragestellung ableiten:

Fragestellung 1:

Verändern Personen ihre mentalen Modelle in Abhängigkeit der jeweiligen Lernaktivitäten?

Es wird davon ausgegangen, dass die entwickelte multimediale Lernumgebung (Kapitel 7.1.1) die Konstruktion mentaler Modelle nachhaltig beeinflusst. Für die Analyse der lernabhängigen Veränderung individueller MODELLE werden die Diagnoseinstrumente *Oberflächen-Struktur-Vergleich* und *SMD-Technologie* herangezogen. Somit ergeben sich folgende spezifische Fragestellungen:

Fragestellung 1.1: Lassen sich lernabhängige Veränderungen von MODELLEN mittels dem Oberflächen-Struktur-Vergleich diagnostizieren?

Fragestellung 1.2: Lassen sich lernabhängige Veränderungen von MODELLEN mittels der SMD-Technologie diagnostizieren?

Aus den theoretischen Annahmen der *SMD-Technologie* (Kapitel 4) ist zu entnehmen, dass sich MODELLE hinsichtlich deren *relationalen* (Surface-Structure), *strukturellen* (Matching-Structure) und *semantisch-strukturellen* (Deep-Structure) Eigenschaften lernabhängig verändern. Es werden folgende Hypothesen zugrunde gelegt:

$H_{0[1.1]}$: *Die Anzahl der Propositionen (Surface-Structure) der MODELLE verändern sich während des Lernprozesses nicht.*

$H_{1[1.1]}$: *Die Anzahl der Propositionen (Surface-Structure) der MODELLE verändern sich während des Lernprozesses.*

$H_{0[1.2]}$: *Der Durchmesser (Matching-Structure) der MODELLE verändert sich während des Lernprozesses nicht.*

$H_{1[1.2]}$: *Der Durchmesser (Matching-Structure) der MODELLE verändert sich während des Lernprozesses.*

$H_{0[1.3]}$: *Die Ähnlichkeit der MODELLE mit einem Referenzmodell (Deep-Structure) verändert sich während des Lernprozesses nicht.*

$H_{1[1.3]}$: *Die Ähnlichkeit der MODELLE mit einem Referenzmodell (Deep-Structure) verändert sich während des Lernprozesses.*

6.2 PÄDAGOGISCHE INTERVENTIONEN

Seel et al. (2000) nehmen an, dass die Konstruktion mentaler Modelle maßgeblich von den dargebotenen Lehrinterventionen einer Lernumgebung abhängt. Aus pädagogischer Sicht schließt sich folglich die zweite übergeordnete Frage an, ob spezielle Interventionen die Konstruktion von mentalen Modellen beeinflussen.

Fragestellung 2:

Welchen Einfluss haben unterschiedliche Lehrinterventionen auf die Konstruktion mentaler Modelle?

Neben mehreren Messzeitpunkten sieht das allgemeine Untersuchungsdesign (vgl. Kapitel 7.1.2) unterschiedliche Treatments (*scaffolding-based learning, self-guided learning, Kontrollgruppe*) vor. Im Hinblick auf die Diagnoseinstrumente *Oberflächen-Struktur-Vergleich* und *SMD-Technologie* ergeben sich folgende spezifische Fragestellungen:

Fragestellung 2.1: Lassen sich Unterschiede zwischen den Treatments (scaffolding-based learning, self-guided learning, Kontrollgruppe) mittels dem Oberflächen-Struktur-Vergleich diagnostizieren?

Fragestellung 2.2: Lassen sich Unterschiede zwischen den Treatments (scaffolding-based learning, self-guided learning, Kontrollgruppe) mittels der SMD-Technologie diagnostizieren?

Empirische Studien (vgl. Anzai & Yokoyama, 1984; Mayer, 1989; Seel & Dinter, 1995; Al-Diban, 2002) zeigen, dass die Konstruktion mentaler Modelle

durch instruktionale Maßnahmen beeinflusst werden kann. Es wird weiter angenommen, dass eine speziell für das modellzentrierte Lernen entwickelte multimediale Lernumgebung mit prozessorientierten Lernhilfen (*Treatments*) die Konstruktion von mentalen Modellen unterstützt (vgl. Seel, 2001). Aus diesen Annahmen werden folgende Hypothesen formuliert:

$H_{0[2,1]}$: Die lernabhängige Veränderung der Anzahl der Propositionen (*Surface-Structure*) der MODELLE unterscheiden sich nicht zwischen den Untersuchungsgruppen (*scaffolding-based learning*, *self-guided learning*, Kontrollgruppe).

$H_{1[2,1]}$: Die lernabhängige Veränderung der Anzahl der Propositionen (*Surface-Structure*) der MODELLE unterscheiden sich zwischen den Untersuchungsgruppen (*scaffolding-based learning*, *self-guided learning*, Kontrollgruppe).

$H_{0[2,2]}$: Der lernabhängige Veränderung des Durchmessers (*Matching-Structure*) der MODELLE unterscheidet sich nicht zwischen den Untersuchungsgruppen (*scaffolding-based learning*, *self-guided learning*, Kontrollgruppe).

$H_{1[2,2]}$: Der lernabhängige Veränderung des Durchmessers (*Matching-Structure*) der MODELLE unterscheidet sich zwischen den Untersuchungsgruppen (*scaffolding-based learning*, *self-guided learning*, Kontrollgruppe).

$H_{0[2,3]}$: Die lernabhängige Veränderung der Ähnlichkeit der MODELLE mit einem Referenzmodell (*Deep-Structure*) unterscheidet sich nicht zwischen den Untersuchungsgruppen (*scaffolding-based learning*, *self-guided learning*, Kontrollgruppe).

$H_{1[2,3]}$: Die lernabhängige Veränderung der Ähnlichkeit der MODELLE mit einem Referenzmodell (*Deep-Structure*) unterscheidet sich zwischen den Untersuchungsgruppen (*scaffolding-based learning*, *self-guided learning*, Kontrollgruppe).

6.3 EFFEKTE VON LERNERVORAUSETZUNGEN AUF DIE KONSTRUKTION MENTALER MODELLE

Lernumgebungen und darin implementierte pädagogische Interventionen können präzise platziert und unterstützend in Bezug auf die mentale Modellbildung eingesetzt werden. Es ist jedoch zu vermuten, dass auch Lernermerkmale wie z. B. *bereichsspezifisches Vorwissen*, *Lernstrategien*, *allgemeine kognitive Fähigkeiten* und *motivationale Merkmale*²⁶ die Konstruktion von mentalen Modellen beeinflussen. Aus diesen Überlegungen ergibt sich die dritte übergeordnete Fragestellung:

Fragestellung 3:

Welchen Einfluss haben unterschiedliche Lernervoraussetzungen auf die Konstruktion mentaler Modelle?

Die Lernervoraussetzungen *Lernstrategien* und *allgemeine kognitive Fähigkeiten* wurden von Couné et al. (2003, 2004) detailliert beleuchtet. Den empirischen Analysen der Autoren ist zu entnehmen, dass die angegebenen Lernervoraussetzungen keinen signifikanten Effekt auf die Qualität der mentalen Modellbildung haben. Auch Pirnay-Dummer et al. (2006) berichten, dass die Lernstrategie-Skalen *Organisieren*, *Elaborieren*, *kritisches Prüfen* und *Metakognitive Strategien* keinen signifikanten Effekt auf die Qualität der mentalen Modellbildung haben. Dennoch soll der Einfluss von *bereichsspezifischem Vorwissen* auf die Konstruktion der MODELLE geprüft werden. Im Hinblick auf das Diagnoseinstrument *SMD-Technologie* ergibt sich folgende spezifische Fragestellung:

Fragestellung 3.1: Welchen Einfluss hat das bereichsspezifische Vorwissen auf die lernabhängige Veränderung von mentalen Modellen, diagnostiziert mit der SMD-Technologie?

Aus der oben formulierten Fragestellung leiten sich folgenden Hypothesen bezüglich der Diagnose mittels *SMD-Technologie* ab:

²⁶ Über den Einfluss motivationaler Merkmale auf die mentale Modellbildung berichtet Couné (2006).

$H_{0[3,1]}$: Das bereichsspezifisches Vorwissen hat keinen Einfluss auf die lernabhängige Veränderung der Anzahl der Propositionen (Surface-Structure) der konstruierten MODELLE.

$H_{1[3,1]}$: Das bereichsspezifisches Vorwissen beeinflusst die lernabhängige Veränderung der Anzahl der Propositionen (Surface-Structure) der konstruierten MODELLE.

$H_{0[3,2]}$: Das bereichsspezifisches Vorwissen hat keinen Einfluss auf die lernabhängige Veränderung des Durchmessers (Matching-Structure) der konstruierten MODELLE.

$H_{1[3,2]}$: Das bereichsspezifisches Vorwissen beeinflusst die lernabhängige Veränderung des Durchmessers (Matching-Structure) der konstruierten MODELLE.

$H_{0[3,3]}$: Das bereichsspezifisches Vorwissen steht in keinem Zusammenhang mit der lernabhängigen Veränderung der Ähnlichkeit zu einem Referenzmodell (Deep-Structure) der konstruierten MODELLE.

$H_{1[3,3]}$: Das bereichsspezifisches Vorwissen steht in Zusammenhang mit der lernabhängigen Veränderung der Ähnlichkeit zu einem Referenzmodell (Deep-Structure) der konstruierten MODELLE.

7 METHODE

“No amount of experimentation can ever prove me right; a single experiment can prove me wrong.” (Albert Einstein, dt.-amerik. Physiker, 1879-1955)

Die im Folgenden dargestellten empirischen Untersuchungen wurden vor dem Hintergrund der *Design-Experimente* (vgl. Brown, 1992) realisiert. Neben der technologischen und didaktischen Entwicklung einer multimedialen Lernumgebung wurden drei experimentelle Studien zu unterschiedlichen Curriculeinheiten durchgeführt. Somit soll Browns (1992) Forderung nach einem Transfer von experimentellen Laborsituationen in den pädagogischen Kontext des *alltäglichen* Klassenzimmers Folge getragen werden.

Eine theoriegeleitet-empirische Untersuchung von mentalen Modellen muss jedoch bestimmte Charakteristika des zu messenden Konstruktes beachten (vgl. Kapitel 2 bis 4). So kann aus methodologischer Sicht nicht das mentale Modell, sondern nur die externalisierte Repräsentation eines mentalen Modells untersucht werden. Des Weiteren bleibt anzumerken, dass mentale Modelle *ad hoc* Konstruktionen sind, die nur in bestimmten Situationen konstruiert werden. Dies schränkt die quantitative Analyse in dem Sinne ein, dass lediglich Bruchstücke von externalisierten Repräsentationen mentaler Modelle diagnostiziert werden können.

Im Rahmen des DFG-Projektes „Modellbegründetes Lernen und Lehren. Multimedialen Gelegenheiten zum Nachdenken“²⁷ konnte ein experimentelles Untersuchungsdesign entwickelt werden, um eine theorieangemessene Diagnose mentaler Modelle zu ermöglichen. Es wurde ein längsschnittliches Untersuchungsdesign realisiert, um die lernabhängigen Prozesse bei der Konstruktion, Modifikation und Reorganisation von Modellen zu diagnostizieren. Die Realisierung der Lernumgebung (7.1.1), der Untersuchungsplan (7.1.2) und die Variablen der Untersuchung (7.1.3) beleuchten die Anordnung der drei durchgeführten experimentellen Studien. Die Stichproben werden in Kapitel 7.2 be-

²⁷ Das an der Albert-Ludwigs-Universität Freiburg durch das Institut für Erziehungswissenschaft durchgeführte Projekt wurde durch die Deutsche Forschungsgemeinschaft unter dem Aktenzeichen SE399/8-1 gefördert.

schrieben. Schließlich zeigen die Datenerhebung (Kapitel 7.3), methodologische Anmerkungen (Kapitel 7.4) und die statistischen Auswertungsverfahren (Kapitel 7.5) das methodische Vorgehen in dieser Arbeit auf.

7.1 ANORDNUNG DER EXPERIMENTELLEN STUDIEN

Um in den experimentellen Studien eine Konstruktion, Revision und Modifikation von Modellen auf Seiten der Probanden zu ermöglichen, wurden in Anlehnung an die Theorie der mentalen Modelle folgende Designprinzipien zugrunde gelegt (vgl. Seel, 2001; Couné et al., 2003).

- Die Lernumgebung solle eine *Problemstellung* enthalten, die in einen authentischen Kontext eingebunden ist.
- Die Lernumgebung solle Informationen in *unterschiedlichen Repräsentationsformaten* und in *non-linearer Anordnung* darbieten. Auf Grund der nicht linearen Struktur sollen die Probanden dazu angeregt werden, verschiedene Informations- und Kommunikationstechnologien aktiv zu nutzen und als Mittel für die Repräsentation und Kommunikation des eigenen Wissens anzuwenden.
- Die Lernumgebung solle *Analogiebildung* anregen. Analogiebildung stellt eine wesentliche Funktion der mentalen Modelle dar (vg. Seel, 1991), wobei empirische Untersuchungen zeigen, dass Probanden gerade der Analogieschluss zwischen verschiedenen Gegenstandsbereichen beim Problemlösen oft Schwierigkeiten bereitet (vgl. Ross, 1987; 1989).
- Die Lernumgebung solle *prozessorientierte Lernhilfen* darbieten. Probanden sollen durch die unterstützenden Maßnahmen an die Problemstellung herangeführt und zum Reflektieren des eigenen Lernprozesses angeregt werden (vgl. Coleman, 1998).
- Die Lernumgebung solle den Probanden die Möglichkeit bieten, ihr *strukturelles und semantisches Wissen zu explizieren*.

7.1.1 Realisierung der Lernumgebung

Ausgehend von den oben genannten Designprinzipien wurde für die drei experimentellen Studien eine multimediale Lernumgebung entwickelt und implementiert. Zentrale Forderung bei der Entwicklung der Lernumgebung war,

dass diese „problemorientiertes und entdeckendes Lernen fördern und signifikante Lernergebnisse, wie z. B. tiefes Verständnis, konzeptuelle Veränderungen und einen Transfer von Wissen und Fertigkeiten, erzielen“ könne (Seel, 2001, p. 8).

Der modulare Aufbau der Lernumgebung (vgl. Abbildung 17) ermöglicht es, einzelne Inhaltsbereiche auszutauschen. So konnten für die drei experimentellen Studien unterschiedliche Inhaltsdomänen von den Probanden bearbeitet werden, ohne dass eine umfangreiche Neuprogrammierung der multimedialen Lernumgebung erfolgen musste (vgl. Couné et al., 2003).

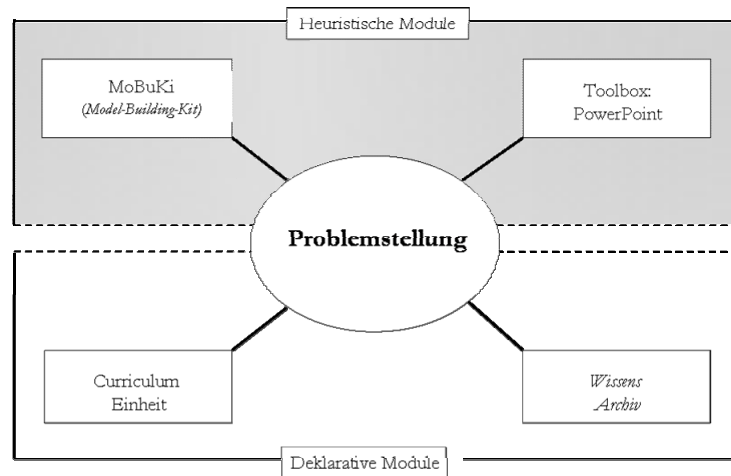


Abbildung 17: Allgemeines Design der multimedialen Lernumgebung

Der non-lineare Aufbau der multimedialen Lernumgebung ermöglicht es den Probanden frei zu navigieren und den individuellen Lernweg selbst zu bestimmen (vgl. Dummer & Ifenthaler, 2005). Die dargebotenen Informationen weisen zusätzlich unterschiedliche Schwierigkeitsgrade auf, was mit Annahmen der Leistungsmotivation korrespondiert (vgl. Heckhausen, 1963, 1989; Atkinson, 1964; Weiner, 1976). So sollen *schwächere* Probanden nicht überfordert und *stärkere* Probanden nicht unterfordert werden. Die Navigation innerhalb der multimedialen Lernumgebung ist so angelegt, dass jederzeit zu allen Modulen gewechselt werden kann. Zusätzlich wurden *dynamische, prozessorientierte Lernhilfen* innerhalb der Lernumgebung implementiert. Diese so genannte *Scaffolding-Komponente* stellt den Probanden an beliebigen Stellen der Lernumgebung zusätzliche Informationen in Form von Aufgaben und Informationstexten

ten bereit. Dadurch soll der individuelle Lernprozess strukturiert und reflektiert, die Problemstellung fokussiert und Hypothesen bezüglich der Fragestellung generiert und überprüft werden (vgl. Couné et al., 2003).

Die zu bearbeitende *Problemstellung* in Form von statischem Text und Bildmaterialien stellt den Ausgangspunkt der multimedialen Lernumgebung dar. Von dort navigieren die Probanden in den *deklarativen* und *heuristischen Modulen* der Lernumgebung, um Informationen zur Lösung des Problems zu bearbeiten.

Die deklarativen Module *Curriculum Einheit* und *Wissens Archiv* beinhalten die für die Problemlösung relevanten Informationen. Im *Wissens Archiv* werden den Probanden Informationen auf vielfältige Art und Weise und in unterschiedlichen Repräsentationsformaten (Text, Audio, Abbildungen, Bilder) dargeboten. Die *Curriculum Einheit* stellt das deklarative Hauptmodul zur jeweiligen Inhaltsdomäne²⁸ dar. Bei der Erstellung und Auswahl der Inhalte einer *Curriculum Einheit* wurde stets darauf geachtet, dass alle notwendigen deklarativen Informationen für die Problemstellung enthalten sind und kein konzeptuelles Modell (vgl. Kapitel 2.1.2) bezüglich der Problemstellung dargestellt wird.

Hauptbestandteil des heuristischen Moduls ist der so genannte *Modellbaukasten* MoBuKi (**ModelBuildingKit**). Hier werden die Probanden angeleitet, was unter der Konzeption „Modell“ zu verstehen ist, wie individuelle Modelle konstruiert werden und welchen Stellenwert die Analogiebildung bei der Modellkonstruktion einnimmt (vgl. Couné et al., 2004). Innerhalb des *Modellbaukastens* können die Probanden zwischen unterschiedlichen Abstraktionsniveaus²⁹ wählen. Auf den drei Abstraktionsniveaus erhalten die Probanden aus unterschiedlichen Perspektiven und durch vielfältige Repräsentationsformate (Grafiken, Animationen, Texte) Informationen bezüglich Hintergründen, Nutzen und Verwendung von Modellen.

²⁸ Die Inhaltsdomäne der ersten Studie wurde aus dem Bereich *Geologie* gewählt. Für die Studien 2 und 3 wurde die Inhaltsdomäne *Geophysik* gewählt.

²⁹ Die drei Abstraktionsniveaus des MoBuKi sind: (1) *Beispiel* (anschauliche Information zur Modell- und Analogiebildung), (2) *Vorgehen* (Information zur schrittweisen Konstruktion von Modellen) und (3) *Wissen* (elaborierte Informationen zu Modellen).

Weiteres heuristisches Modul ist die *Toolbox*, welche durch MS PowerPoint (Microsoft Corporation, 2000) ergänzt wurde (vgl. Seel, 2003c). Mit Hilfe der *Toolbox* externalisieren die Probanden ihre mentalen Modelle mittels eines vorab kommunizierten Notationssystems, bestehend aus Knoten und Relationen (vgl. Kapitel 4). Diese MODELLE wurden an definierten Messzeitpunkten gespeichert und dienen als Datengrundlage für die empirische Analyse der lernabhängigen Veränderung. Die einzelnen Phasen der drei experimentellen Studien werden im Untersuchungsplan vorgestellt.

7.1.2 Untersuchungsplan

Wie bereits beschrieben, wurden die drei experimentellen Studien dieser Arbeit im Rahmen des DFG-Projektes „Modellbegründetes Lernen und Lehren. Multimedialen Gelegenheiten zum Nachdenken“ realisiert. Der allgemeine Untersuchungsplan illustriert das längsschnittliche, quasi-experimentelle Design der drei aufeinander folgend durchgeführten Studien (vgl. Abbildung 18).

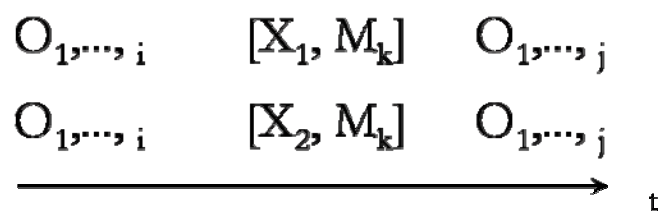


Abbildung 18: Allgemeines Versuchsdesign der drei experimentellen Studien

Dabei bedeuten:

$O_{1,\dots,i}$	Messungen der unabhängigen Variablen
X_1	Treatment <i>scaffolding-based learning</i> (prozessorientierte Lernhilfen)
X_2	Treatment <i>self-guided learning</i> (keine expliziten Lernhilfen)
M_k	Messung der konstruierten MODELLE
$O_{1,\dots,j}$	Messungen der abhängigen Variablen nach Behandlungskomplexion

Auf Grund des zeitintensiven Untersuchungsplans konnte keine zufällige Zuordnung der Probanden erfolgen. Die Probanden der drei Stichproben (N

= 106) wurden an Freiburger Realschulen und Gymnasien und an der Albert-Ludwigs-Universität Freiburg akquiriert. Durchgeführt wurden die drei Studien im Multimedialabor³⁰ des Instituts für Erziehungswissenschaft an der Albert-Ludwigs-Universität Freiburg. Die Probanden wurden zu Beginn der Studien per Zufall den experimentellen Bedingungen *scaffolding-based learning* und *self-guided learning* zugeordnet. Außerdem wurden die Probanden zufällig einer *individuellen* oder *kollaborativen Lernform* zugewiesen. Die experimentelle Bedingung *kollaborative Lernform* wurde im Rahmen der Analysen des DFG-Projektes weiterverfolgt (vgl. Couné et al., 2003, 2004). In Übereinstimmung mit der Theorie der mentalen Modelle (vgl. Seel, 1991), nach der mentale Modelle *subjektive ad-hoc Konstruktionen* sind, wird eine Analyse der *kollaborativen Lernform* in dieser Arbeit nicht weiterverfolgt. Daher werden ausschließlich die grau hinterlegten Zellen des 2x2-faktoriellen quasi-experimentellen Designs dargestellt und analysiert (vgl. Abbildung 19).

	Scaffolding-based learning X_1	Self-guided learning X_2
Individuelles Lernen Y_1		
Kollaboratives Lernen Y_2		

Abbildung 19: 2x2-faktorielles Design der experimentellen Studien

Die übergeordnete Fragestellung dieser Arbeit nach einer Diagnose lernabhängiger Veränderung mentaler Modelle erfordert einen längsschnittlichen Untersuchungsplan. Der Untersuchungsplan sieht daher pro Studie acht Mess-

³⁰ Das Multimedialabor des Instituts für Erziehungswissenschaft verfügt über speziell ausgestattete Rechner, die in Form eines Netzwerkes verbunden sind. Die technische Ausstattung der Rechner gestattet es, umfangreiche Anwendungen und verschiedene Repräsentationsformate (Text, Audio, Video, Shockwave, etc.) zu implementieren. Die Netzwerktechnik und die implementierte multimediale Lernumgebung ermöglichen es, individuelle Lernpfade der Probanden zu speichern und zu analysieren (vgl. Dummer & Ifenthaler, 2005).

zeitpunkte (MZP₀ bis MZP₇) vor. Die Besonderheiten der drei experimentellen Studien werden in den folgenden Abschnitten (7.1.2.1 bis 7.1.2.3) detailliert betrachtet.

7.1.2.1 Experimentelle Studie 1

Als Inhaltsdomäne der ersten Studie wurde aus dem Bereich *Geologie* das Thema *Ökosystem Wald* gewählt. Zum MZP₀ teilten die Versuchsleiter die Probanden per Zufall den experimentellen Bedingungen zu. Nach der Bearbeitung des *allgemeinen Fragebogens* (demographische Daten) wurden drei Skalen des *Berliner Intelligenz Struktur Tests* (Jäger et al., 1997) erhoben³¹. Zum MZP₁ bearbeiteten die Probanden die Form A des *Fragebogens zur Selbstwirksamkeit und Interesse* (vgl. Couné, 2006), anschließend den *LIST-Fragebogen* (Wild & Schiefele, 1994) und die Form A des *deklarativen Wissenstests* (18 Multiple-Choice-Fragen zum Thema Biologie). Anschließend wurde ein standardisiertes Training zur Notation und Konstruktion von MODELLEN durchgeführt. Zum Abschluss wurden die Probanden mit der *Problemstellung* konfrontiert und sollten darauf hin innerhalb 15 Minuten ein erstes MODELL (PRÄKONZEPTION) konstruieren. Die Probanden sollten folgende *Problemstellung* bearbeiten:

Erkläre den Material- und Energiefluss im Ökosystem Wald! Wer ist beteiligt? Wie hängen die Dinge zusammen?

An den folgenden fünf MZPen (MZP₂₋₆) arbeiteten die Probanden selbstgesteuert jeweils eine Stunde mit der multimedialen Lernumgebung. Die zu jedem MZP konstruierten MODELLE wurden durch den Versuchsleiter gespeichert. Als Abschluss jedes MZP wurde ein *kurzer Test zur Selbstwirksamkeit und Interesse* bearbeitet (vgl. Couné, 2006). Zum MZP₇ präsentierten die Probanden den anderen anwesenden Probanden das letzte MODELL der individuellen Problemlösung. Danach wurde die Form B des *deklarativen Wissenstests* durchgeführt. Ohne multimediale Lernumgebung sollten die Probanden innerhalb von 15 Minuten ein letztes MODELL zur *Problemstellung* konstruieren. Abschließend wurde die Form B des *Fragebogens zur Selbstwirksamkeit und Interesse* bearbeitet.

³¹ Eine ausführliche Beschreibung der verwendeten Instrumente erfolgt in Kapitel 7.4.

Probanden mit der experimentellen Bedingung *scaffolding-based learning* erhielten während der Arbeit mit dem multimedialen Lernprogramm prozessorientierte Lernhilfen. Diese wurden dynamisch durch das System generiert und fokussierten auf die Problemanalyse, die Zielformulierung und –überprüfung, sowie die Hypothesengenerierung. Außerdem erhielten die Probanden mit der experimentellen Bedingung *scaffolding-based learning* zum MZP₄ ein detailliertes schriftliches Feedback zum Modell des MZP₃, in Form von Fragen, welche auf Verbesserungsmöglichkeiten hinwiesen. Im Gegensatz dazu, erhielten die Probanden mit der experimentellen Bedingung *self-guided learning* ein allgemeines Feedback zum Modell des MZP₃ (vgl. Abbildung 20).

	Scaffolding-based learning X ₁	Self-guided learning X ₂
MZP ₀	1. Allg. Fragebogen 2. BIS-Test	1. Allg. Fragebogen 2. BIS-Test
MZP ₁	1. SW-Fragebogen A 2. LIST-Fragebogen 3. deklarativer Wissenstest A 4. Training zur Notation und Konstruktion von Modellen 5. Erhebung Modell ₁	1. SW-Fragebogen A 2. LIST-Fragebogen 3. deklarativer Wissenstest A 4. Training zur Notation und Konstruktion von Modellen 5. Erhebung Modell ₁
MZP ₂	1. Lernumgebung mit prozessorientierten Lernhilfen 2. Erhebung Modell ₂ 3. SW-Kurzfragebogen	1. Lernumgebung ohne prozessorientierte Lernhilfen 2. Erhebung Modell ₂ 3. SW-Kurzfragebogen
MZP ₃	1. Lernumgebung mit prozessorientierten Lernhilfen 2. Erhebung Modell ₃ 3. SW-Kurzfragebogen	1. Lernumgebung ohne prozessorientierte Lernhilfen 2. Erhebung Modell ₃ 3. SW-Kurzfragebogen
MZP ₄	1. Lernumgebung mit prozessorientierten Lernhilfen 2. Detailliertes Modell-Feedback 3. Erhebung Modell ₄ 4. SW-Kurzfragebogen	1. Lernumgebung ohne prozessorientierte Lernhilfen 2. Allgemeines Modell-Feedback 3. Erhebung Modell ₄ 4. SW-Kurzfragebogen
MZP ₅	1. Lernumgebung mit prozessorientierten Lernhilfen 2. Erhebung Modell ₅ 3. SW-Kurzfragebogen	1. Lernumgebung ohne prozessorientierte Lernhilfen 2. Erhebung Modell ₅ 3. SW-Kurzfragebogen
MZP ₆	1. Lernumgebung mit prozessorientierten Lernhilfen 2. Erhebung Modell ₆ 3. SW-Kurzfragebogen	1. Lernumgebung ohne prozessorientierten Lernhilfen 2. Erhebung Modell ₆ 3. SW-Kurzfragebogen
MZP ₇	1. Präsentation der individuellen Lösung 2. deklarativer Wissenstest B 3. Erhebung Modell ₇ 4. SW-Fragebogen B	1. Präsentation der individuellen Lösung 2. deklarativer Wissenstest B 3. Erhebung Modell ₇ 4. SW-Fragebogen B

Abbildung 20: Untersuchungsplan der Studie 1

7.1.2.2 Experimentelle Studie 2

Als Inhaltsdomäne der zweiten Studie wurde aus dem Bereich *Geophysik* das Thema *Plattentektonik* gewählt. Analog zur ersten Studie teilten die Versuchsleiter zum MZP₀ die Probanden per Zufall den experimentellen Bedingungen zu und erhoben die demographischen Daten per *allgemeinem Fragebogen*. Anschließend wurden drei Skalen des *Berliner Intelligenz Struktur Tests* (Jäger et al., 1997) durchgeführt. Zum MZP₁ bearbeiteten die Probanden die Form A des *Fragebogens zur Selbstwirksamkeit und Interesse* und des *deklarativen Wissenstest* und den *LIST-Fragebogen*. Nach dem Training zur Notation und Konstruktion von MODELLEN wurden die Probanden mit der *Problemstellung* konfrontiert und sollten darauf hin innerhalb 15 Minuten ein erstes MODELL (PRÄKONZEPTION) konstruieren. Für die zweite Studie sollte folgende *Problemstellung* bearbeitet werden:

„Was wäre denn, wenn es die Plattentektonik nicht mehr gäbe? Welche Konsequenzen hätte dies auf wen und was und warum?“

An den folgenden fünf Messzeitpunkten (MZP₂₋₆) arbeiteten die Probanden selbstgesteuert jeweils eine Stunde mit der multimedialen Lernumgebung. Die zu jedem MZP konstruierten MODELLE wurden durch den Versuchsleiter gespeichert. Als Abschluss jedes MZP wurde ein *kurzer Test zur Selbstwirksamkeit und Interesse* bearbeitet (vgl. Couné, 2006). Zum MZP₇ präsentierten die Probanden den anderen Probanden das letzte MODELL der individuellen Problemlösung. Danach wurde die Form B des *deklarativen Wissenstests* durchgeführt. Ohne multimediale Lernumgebung sollten die Probanden innerhalb von 15 Minuten ein letztes MODELL zur *Problemstellung* konstruieren. Abschließend wurde die Form B des *Fragebogens zur Selbstwirksamkeit und Interesse* bearbeitet.

Probanden mit der experimentellen Bedingung *scaffolding-based learning* erhielten während der Arbeit mit dem multimedialen Lernprogramm prozessorientierte Lernhilfen. Diese wurden dynamisch durch das System generiert und fokussierten auf die Problemanalyse, die Zielformulierung und –überprüfung, sowie die Hypothesengenerierung. Außerdem erhielten die Probanden mit der experimentellen Bedingung *scaffolding-based learning* zum MZP₅ ein detailliertes, schriftliches Feedback zum Modell des MZP₄, in Form von Fragen, welche auf Verbesserungsmöglichkeiten hinwiesen.

Die Probanden mit der experimentellen Bedingung *self-guided learning* erhielten kein Feedback zum MZP₅ (vgl. Abbildung 21).

	Scaffolding-based learning X ₁	Self-guided learning X ₂
MZP ₀	1. Allg. Fragebogen 2. BIS-Test	1. Allg. Fragebogen 2. BIS-Test
MZP ₁	1. SW-Fragebogen A 2. deklarativer Wissenstest A 3. LIST-Fragebogen 4. Training zur Notation und Konstruktion von Modellen 5. Erhebung Modell ₁	1. SW-Fragebogen A 2. deklarativer Wissenstest A 3. LIST-Fragebogen 4. Training zur Notation und Konstruktion von Modellen 5. Erhebung Modell ₁
MZP ₂	1. Lernumgebung mit prozessorientierten Lernhilfen 2. Erhebung Modell ₂ 3. SW-Kurzfragebogen	1. Lernumgebung ohne prozessorientierte Lernhilfen 2. Erhebung Modell ₂ 3. SW-Kurzfragebogen
MZP ₃	1. Lernumgebung mit prozessorientierten Lernhilfen 2. Erhebung Modell ₃ 3. SW-Kurzfragebogen	1. Lernumgebung ohne prozessorientierte Lernhilfen 2. Erhebung Modell ₃ 3. SW-Kurzfragebogen
MZP ₄	1. Lernumgebung mit prozessorientierten Lernhilfen 2. Detailliertes Modell-Feedback 3. Erhebung Modell ₄ 4. SW-Kurzfragebogen	1. Lernumgebung ohne prozessorientierte Lernhilfen 2. Erhebung Modell ₄ 3. SW-Kurzfragebogen
MZP ₅	1. Lernumgebung mit prozessorientierten Lernhilfen 2. Erhebung Modell ₅ 3. SW-Kurzfragebogen	1. Lernumgebung mit prozessorientierten Lernhilfen 2. Erhebung Modell ₅ 3. SW-Kurzfragebogen
MZP ₆	1. Lernumgebung mit prozessorientierten Lernhilfen 2. Erhebung Modell ₆ 3. SW-Kurzfragebogen	1. Lernumgebung ohne prozessorientierte Lernhilfen 2. Erhebung Modell ₆ 3. SW-Kurzfragebogen
MZP ₇	1. Präsentation der individuellen Lösung 2. deklarativer Wissenstest B 3. Erhebung Modell ₇ 4. SW-Fragebogen B	1. Präsentation der individuellen Lösung 2. deklarativer Wissenstest B 3. Erhebung Modell ₇ 4. SW-Fragebogen B

Abbildung 21: Untersuchungsplan der Studie 2

7.1.2.3 Experimentelle Studie 3

Analog zur zweiten Inhaltsdomäne wurde für die dritte Studie das Thema *Plattentektonik* aus dem Bereich *Geophysik* gewählt. Zusätzlich zu den experimentellen Bedingungen *scaffolding-based learning* und *self-guided learning* wurde eine *Kontrollgruppe* hinzugefügt. Die Probanden der *Kontrollgruppe* arbeiteten mit einer multimedialen Lernumgebung zum Thema Gentechnik. Die letzten 15 Minuten eines MZP konstruierten die Probanden der *Kontrollgruppe* ein MODELL zur

Problemstellung (identisch mit der *Problemstellung* der beiden anderen experimentellen Variationen) der dritten Studie. Außerdem wurden die experimentellen Bedingungen *scaffolding-based learning* und *self-guided learning* variiert. Die Ergebnisse dieser Variation werden von Hanke (2006) dargestellt und diskutiert und finden in dieser Arbeit keine nähere Betrachtung. Analog zur ersten und zweiten Studie teilten die Versuchsleiter zum MZP₀ die Probanden per Zufall den experimentellen Bedingungen zu und erhoben die demographischen Daten per *allgemeinem Fragebogen*. Anschließend wurden drei Skalen des *Berliner Intelligenz-Struktur Tests* (Jäger et al., 1997) durchgeführt. Zum MZP₁ bearbeiteten die Probanden die Form A des *Fragebogens zur Selbstwirksamkeit und Interesse* und des *deklarativen Wissenstest* und den *LIST-Fragebogen*. Nach dem Training zur Notation und Konstruktion von MODELLEN wurden die Probanden mit der *Problemstellung* konfrontiert und sollten darauf hin innerhalb 15 Minuten ein erstes MODELL (PRÄKONZEPTION) konstruieren.

Für die dritte Studie sollte folgende *Problemstellung* bearbeitet werden:

Was wäre denn, wenn es die Plattentektonik nicht mehr gäbe? Welche Konsequenzen hätte dies auf wen und was und warum?

An den folgenden fünf MZPen (MZP₂₋₆) arbeiteten die Probanden selbstgesteuert jeweils eine Stunde mit der multimedialen Lernumgebung. Die zu jedem MZP konstruierten MODELLE wurden durch den Versuchsleiter gespeichert. Als Abschluss jedes MZP wurde ein *kurzer Test zur Selbstwirksamkeit und Interesse* bearbeitet (vgl. Couné, 2006). Zum MZP₇ präsentierten die Probanden den anderen Probanden das letzte MODELL der individuellen Problemlösung. Danach wurde die Form B des *deklarativen Wissenstests* durchgeführt. Ohne multimediale Lernumgebung sollten die Probanden innerhalb von 15 Minuten ein letztes MODELL zur *Problemstellung* konstruieren. Abschließend wurde die Form B des *Fragebogens zur Selbstwirksamkeit und Interesse* bearbeitet.

Probanden mit der experimentellen Bedingung *scaffolding-based learning* erhielten während der Arbeit mit dem multimedialen Lernprogramm prozessorientierte Lernhilfen. Diese wurden dynamisch durch das System generiert und fokussierten auf die Problemanalyse, die Zielformulierung und –überprüfung, sowie die Hypothesengenerierung. Außerdem erhielten die Probanden mit der experimentellen Bedingung *scaffolding-based learning* zum MZP₅ ein schriftliches detailliertes Feedback zum Modell des MZP₄, in Form von Fragen, welche auf

Verbesserungsmöglichkeiten hinwies. Die Probanden mit der experimentellen Bedingung *self-guided learning* und *Kontrollgruppe* erhielten kein Feedback zum MZP₅ (vgl. Abbildung 22).

	Scaffolding-based leaning X ₁	Self-guided learning X ₂	Kontrollgruppe X ₃
MZP ₀	1. Allg. Fragebogen 2. BIS-Test	1. Allg. Fragebogen 2. BIS-Test	1. Allg. Fragebogen 2. BIS-Test
MZP ₁	1. SW-Fragebogen A 2. Lernstile-Fragebogen 3. deklarativer Wissenstest A 4. LIST-Fragebogen 5. Training zur Notation und Konstruktion von Modellen 6. Erhebung Modell ₁	1. SW-Fragebogen A 2. Lernstile-Fragebogen 3. deklarativer Wissenstest A 4. LIST-Fragebogen 5. Training zur Notation und Konstruktion von Modellen 6. Erhebung Modell ₁	1. SW-Fragebogen A 2. Lernstile-Fragebogen 3. deklarativer Wissenstest A 4. LIST-Fragebogen 5. Training zur Notation und Konstruktion von Modellen 6. Erhebung Modell ₁
MZP ₂	1. Lernumgebung mit prozessorientierten Lernhilfen 2. Erhebung Modell ₂ 3. SW-Kurzfragebogen	1. Lernumgebung ohne prozessorientierte Lernhilfen 2. Erhebung Modell ₂ 3. SW-Kurzfragebogen	1. Lernumgebung zu anderem Inhaltsbereich 2. Erhebung Modell ₂ 3. SW-Kurzfragebogen
MZP ₃	1. Lernumgebung mit prozessorientierten Lernhilfen 2. Erhebung Modell ₃ 3. SW-Kurzfragebogen	1. Lernumgebung ohne prozessorientierte Lernhilfen 2. Erhebung Modell ₃ 3. SW-Kurzfragebogen	1. Lernumgebung zu anderem Inhaltsbereich 2. Erhebung Modell ₃ 3. SW-Kurzfragebogen
MZP ₄	1. Lernumgebung mit prozessorientierten Lernhilfen 2. Detailliertes Modell-Feedback 3. Erhebung Modell ₄ 4. SW-Kurzfragebogen	1. Lernumgebung ohne prozessorientierte Lernhilfen 2. Erhebung Modell ₄ 3. SW-Kurzfragebogen	1. Lernumgebung zu anderem Inhaltsbereich 2. Erhebung Modell ₄ 3. SW-Kurzfragebogen
MZP ₅	1. Lernumgebung mit prozessorientierten Lernhilfen 2. Erhebung Modell ₅ 3. SW-Kurzfragebogen	1. Lernumgebung mit prozessorientierten Lernhilfen 2. Erhebung Modell ₅ 3. SW-Kurzfragebogen	1. Lernumgebung zu anderem Inhaltsbereich 2. Erhebung Modell ₅ 3. SW-Kurzfragebogen
MZP ₆	1. Lernumgebung mit prozessorientierten Lernhilfen 2. Erhebung Modell ₆ 3. SW-Kurzfragebogen	1. Lernumgebung ohne prozessorientierte Lernhilfen 2. Erhebung Modell ₆ 3. SW-Kurzfragebogen	1. Lernumgebung zu anderem Inhaltsbereich 2. Erhebung Modell ₆ 3. SW-Kurzfragebogen
MZP ₇	1. Präsentation der individuellen Lösung 2. deklarativer Wissenstest B 3. Erhebung Modell ₇	1. Präsentation der individuellen Lösung 2. deklarativer Wissenstest B 3. Erhebung Modell ₇	1. Präsentation der individuellen Lösung 2. deklarativer Wissenstest B 3. Erhebung Modell ₇

Abbildung 22: Untersuchungsplan der Studie 3

7.1.3 Variablen der Untersuchung

Nach Kerlinger (1986) ist eine Präzisierung der Untersuchungsvariablen und deren genaue Unterscheidung eine notwendige Bedingung für empirische Analysen. So werden in den folgenden Abschnitten *experimentelle Variablen* (7.1.3.1), *Personenvariablen* (7.1.3.2) und *abhängige Variablen* (7.1.3.3) unterschieden.

7.1.3.1 Experimentelle Variablen

Vor dem Hintergrund der theoretischen Überlegungen (vgl. Kapitel 2 bis 5) und den Fragestellungen (vgl. Kapitel 6) wurden die Probanden in zwei (Studie 1 und 2) bzw. drei (Studie 3) Untersuchungsgruppen eingeteilt.

X ₁	scaffolding-based learning
X ₂	self-guided learning
X ₃	Kontrollgruppe

7.1.3.2 Personenvariablen

Wie die Ausführungen zu mentalen Modellen (Kapitel 2) verdeutlicht haben, wird davon ausgegangen, dass verschiedene Personenvariablen einen differentiellen Einfluss auf das modellzentrierte Lernen ausüben.

7.1.3.2.1 Bereichsspezifisches Vorwissen

Nach Ausubel (1968) beeinflusst das Vorwissen entscheidend kognitive Lernprozesse. Vor diesem Hintergrund wurde in den drei experimentellen Studien das *bereichsspezifische Vorwissen* als Prädiktorvariable erhoben. Das *bereichsspezifische Vorwissen* konnte mittels Multiple-Choice-Fragebogen entsprechend der Inhaltsdomäne erhoben werden (vgl. Kapitel 7.3).

7.1.3.2.2 Allgemeine kognitive Fähigkeiten

Zur Kontrolle der allgemeinen kognitiven Fähigkeiten wurden unter Verwendung des BIS (Jäger et al., 1997) die Skalen *sprachgebundenes Denken*, *figuralbildhaftes Denken* und *Einfallsreichtum* erhoben.

7.1.3.2.3 Lernstrategien

Lernstrategien und metakognitive Fähigkeiten beeinflussen entscheidend Lern- und Problemlöseleistungen (vgl. Schenk, 2003; Vögele, 2004). Im Hinblick auf die weitgehend selbstständigen Lernaktivitäten in den drei experimentellen Studien werden die von den Probanden präferierten Lernstrategien erhoben. Aus dem LIST-Fragebogen (Wild et al., 1992) ergeben sich insgesamt

elf Prädiktorvariablen (in der empirischen Analyse werden lediglich die ersten fünf Variablen berücksichtigt³²):

- **Organisieren**
- **Elaborieren**
- **Kritisches Prüfen**
- **Wiederholen**
- **Metakognitive Strategien**
- Anstrengung
- Aufmerksamkeit
- Zeitmanagement
- Lernumgebung
- Lernen mit Studienkollegen
- Literatur

7.1.3.2.4 Weitere Variablen

Zusätzlich wurden in den drei experimentellen Studien weitere unabhängige Variablen wie *Schulform*, *Bildungsstand* (Schule / Studium), *Geschlecht* und *Alter* in Form eines allgemeinen Fragebogens erhoben.

7.1.3.3 Abhängige Variablen

Ausgehend von der Fragestellung nach einer lernabhängigen Veränderung mentaler Modelle (vgl. Kapitel 6) stehen die individuell konstruierten MODELLE im Mittelpunkt des Untersuchungsinteresses. Unter Verwendung der *SMD-Technologie* (vgl. Kapitel 4) ergeben sich somit drei Kriteriumsvariablen:

- Surface-Structure Index θ
- Matching-Structure Index μ

³² Aus theoretischer Sicht tragen die übrigen Variablen des LIST-Tests nicht zum Erkenntnisgewinn der empirischen Arbeit bei und werden daher nicht weiter berücksichtigt.

- Deep-Structure Index δ

7.1.3.3.1 Surface-Structure Index θ

Der Surface-Structure Index θ bestimmt die *perzipierte Oberflächenstruktur* der konstruierten MODELLE. Dieser wird aus der Summe der verwendeten Propositionen der MODELLE berechnet (vgl. Kapitel 4.4).

7.1.3.3.2 Matching-Structure Index μ

Mit dem Matching-Structure Index μ wird die *abgeleitete Strukturstabilität* der individuellen MODELLE bestimmt. Die strukturellen Veränderungen der MODELLE stellen eine entscheidende Kriteriumsvariable in Bezug auf die lernabhängige Veränderung mentaler Modelle dar.

7.1.3.3.3 Deep-Structure Index δ

Der Deep-Structure Index δ bestimmt die *semantische Strukturcharakteristik* der konstruierten Modelle. Dabei werden die konstruierten Modelle hinsichtlich ihrer semantischen Qualität mit einem Erklärungs- bzw. Expertenmodell geprüft (vgl. Kapitel 4.4).

7.2 STICHPROBEN DER DREI EXPERIMENTELLEN STUDIEN

Wie bereits in Abbildung 12 illustriert, geht nur ein Teil der Gesamtstichprobe des DFG-Projektes in die Analysen dieser Arbeit mit ein. Die einzelnen Stichproben der drei Studien werden in den folgenden Abschnitten (7.2.1 bis 7.2.3) ausführlich dargestellt.

7.2.1 Stichprobe der Studie 1

An der ersten Studie nahmen 26 Schüler (17 weiblich, 9 männlich; $\chi_{(1)}^2 = 2.462$; $p = .117$) der 9. Klassenstufe aus Freiburger Gymnasien teil. Das Durchschnittsalter der Schüler lag bei 15.04 Jahren ($Min = 14$; $Max = 16$; $SD = .599$). Die Schüler erhielten eine Aufwandsentschädigung. Tabelle 2 zeigt die Zutei-

lung der Schüler zu den experimentellen Bedingungen *scaffolding-based learning* und *self-guided learning*.

Tabelle 2: Verteilung der Probanden zu den experimentellen Bedingungen (Studie 1)

	Geschlecht		Gesamt
	männlich	weiblich	
self-guided learning	5	8	13
scaffolding-based learning	4	9	13
Gesamt	9	17	26

7.2.2 Stichprobe der Studie 2

Nachdem es sich in der ersten Studie als schwierig erwies, für mehrere Messzeitpunkte Probanden zu rekrutieren, wurde die Stichprobe in der zweiten Studie mit Schülern der 8., 9. und 10. Klassenstufe aus Freiburger Gymnasien und Realschulen erweitert. Insgesamt 33 Schüler (19 weiblich, 14 männlich; $\chi_{(1)}^2 = .758$; $p = .384$) mit einem Durchschnittsalter von 14.21 Jahren ($Min = 12$; $Max = 16$; $SD = 1.11$) nahmen an der zweiten Studie teil. Auch die Schüler der zweiten Studie erhielten eine Aufwandsentschädigung. Aus Freiburger Gymnasien konnten 15 Schüler und aus Realschulen 18 Schüler gewonnen werden ($\chi_{(1)}^2 = .273$; $p = .602$). Die Zuteilung der Schüler zu den experimentellen Bedingungen *scaffolding-based learning* und *self-guided learning* ist Tabelle 3 zu entnehmen.

Tabelle 3: Verteilung der Probanden zu den experimentellen Bedingungen (Studie 2)

	Geschlecht		Gesamt
	männlich	weiblich	
self-guided learning	6	10	16
scaffolding-based learning	8	9	17
Gesamt	14	19	33

7.2.3 Stichprobe der Studie 3

Für die dritte Studie wurden Studierende des Studiengangs *Bildungsplanung und Instructional Design* gewonnen. Insgesamt nahmen 47 Studierende (34 weiblich, 13 männlich; $\chi_{(1)}^2 = 9.383$; $p = .002$) der Albert-Ludwigs-Universität Freiburg teil. Das Durchschnittsalter der Studierenden lag bei 23.02 Jahren ($Min = 19$; $Max = 32$; $SD = 2.409$). Sie studierten im zweiten bis 24. Semester ($M = 4.77$; $SD = 3.528$). Tabelle 4 zeigt die Zuteilung der Studierenden zu den expe-

rimentellen Bedingungen *scaffolding-based learning*, *self-guided learning* und *Kontrollgruppe*.

Tabelle 4: Verteilung der Probanden zu den experimentellen Bedingungen (Studie 3)

	Geschlecht		Gesamt
	männlich	weiblich	
self-guided learning	3	15	18
scaffolding-based learning	7	12	19
Kontrollgruppe	3	7	10
Gesamt	13	34	47

7.3 INSTRUMENTE

7.3.1 Allgemeiner Fragebogen

Der *allgemeine Fragebogen* wurde zu Beginn der experimentellen Studien von jedem Probanden ausgefüllt. Neben der *Versuchspersonennummer* erhebt der Fragebogen das *Alter* und *Geschlecht* der Probanden, die Nummer der *experimentellen Variation*, *Schultyp bzw. Name der Schule* und ob Interesse an weiteren Studien am Institut für Erziehungswissenschaft besteht. Für Studie 3 wurde an Stelle des Schultypus das aktuelle *Hochschulsemester* und die *Fächerkombination* erhoben (vgl. Couné et al., 2003, 2004).

7.3.2 Deklarativer Wissenstest

Zur jeweiligen Inhaltsdomäne der experimentellen Studien wurde das *deklarative Wissen* der Probanden vor und nach der Lernphase erhoben. Die Multiple-Choice-Fragebögen wurden vor deren Einsatz auf Reliabilität und Validität getestet (vgl. Couné et al., 2003). Der für Studie 1 eingesetzte Wissenstest zum Thema „Ökosystem Wald“ umfasst insgesamt 18 Multiple-Choice-Aufgaben. Vor- und Nachtest sind nahezu identisch, es wurde lediglich die Reihenfolge der Aufgaben variiert (vgl. Couné et al., 2003). Auch der für die Studien 2 und 3 eingesetzte Wissenstest „Geophysik“ wurde vor dem Einsatz auf Reliabilität und Validität getestet (vgl. Couné et al., 2004). Insgesamt enthält der Fragebogen 19 Multiple-Choice-Fragen. Vor- und Nachtest sind nahezu identisch, es wurde lediglich die Reihenfolge der Aufgaben variiert (vgl. Couné et al., 2004).

7.3.3 Berliner Intelligenz-Struktur-Test

Der *Berliner Intelligenz-Struktur-Test* (BIS-Test, Jäger et al., 1997) konnte in zahlreichen Untersuchungen als konstruktvalides Messinstrument getestet werden. Im Hinblick auf die Fragestellung dieser Arbeit wurden die Skalen *sprachgebundenes Denken*, *figural-bildhaftes Denken* und *Einfallsreichtum* erhoben. Da nicht alle Skalen erhoben wurden, kann keine Aussage über die *allgemeine Intelligenz* (Integral aller sieben Skalen) der Probanden gemacht werden.

7.3.4 LIST-Test

Schließlich wurde als Instrument das *Inventar zur Erfassung von Lernstrategien im Studium* (LIST) von Wild & Schiefele (1994) eingesetzt. Die insgesamt 77 Items bilden elf Skalen mit der Bezeichnung *Organisieren*, *Elaborieren*, *Kritisches Prüfen*, *Wiederholen*, *Metakognitive Strategien*, *Anstrengung*, *Aufmerksamkeit*, *Zeitmanagement*, *Lernumgebung*, *Lernen mit Studienkollegen* und *Literatur*. Auch der LIST wurde in zahlreichen empirischen Studien auf Reliabilität und Validität geprüft (vgl. Wild, 2000).

7.3.5 SMD-Technologie

Aufgabe der *SMD-Technologie* ist es, die lernabhängige Veränderung externalisierter MODELLE zu erfassen. Da sich die erhobenen MODELLE als sehr komplex erwiesen und während der drei experimentellen Studien 742 MODELLE konstruiert wurden, wird die *SMD-Technologie* als automatisiertes Diagnoseinstrument eingesetzt (vgl. Kapitel 4.5). Hierfür wurden alle MODELLE von zwei Diagnostikern in Listenform transformiert. Die Ergebnisse der Reliabilitätsprüfung mittels Übereinstimmungskoeffizient Kendall-W ($\tau = .932$) zeigen eine sehr hohe Übereinstimmung der Diagnostiker bei der Transformation der MODELLE in Listenform (vgl. Tabelle 5). Da die Transformation der MODELLE in Listen den einzigen manuellen Schritt der *SMD-Technologie* darstellt und alle weiteren Analyseschritte automatisiert erfolgen, kann dem Instrument folglich eine sehr hohe Zuverlässigkeit zugeschrieben werden.

Tabelle 5: Reliabilität der Transformation der MODELLE in Listenform

	Listentrans- formation
N Diagnostiker	2
Chi-Quadrat	283.296
df	152
Kendalls-W	.932
Asymptotische Signifikanz	.000

Hinsichtlich der Validität der *SMD-Technologie* (siehe auch Kapitel 7.4.4) konnte eine Kreuzvalidierung mit mehreren quantitativen und qualitativen Forschungsmethodologien zur Erfassung von externalisierten Modellen erfolgen. So wurden in einer Studie (vgl. Johnson et al., 2006) die quantitativen Methodologien der *SMD-Technologie* (vgl. Kapitel 4) und MITOCAR (vgl. Pirnay-Dummer, 2006) und die qualitativen Methodologien DEEP (Spector & Koszalka, 2004) und ACSMM (O'Connor & Johnson, 2004) hinsichtlich der Ergebnisse aus einem gemeinsamen Datensatz validiert. Die Methodologien unterscheiden sich hinsichtlich deren Datenaufbereitung, Datenerfassung, Analyseverfahren und Vergleichsmöglichkeiten (vgl. Tabelle 6).

Tabelle 6: Vergleich der vier Methodologien zur Erfassung von externalisierten Modellen (Johnson et al., 2006)

Method	Data Collection	Analysis	Data Conversion	Comparison Fxn
ACSMM	Concept Map	Qualitative with Descriptive Statistics—analysis is done mostly by hand	Structural Decomposition into 3 Key Categories (manual), Structural Re-composition into 1 Representation	Unlimited comparisons, show details relating to concepts
SMD	Concept Map or Natural Language	Quantitative—analysis is calculated using tools	Structural Decomposition into 3 Key Categories (manual)	Unlimited comparisons
MITOCAR	Natural Language	Quantitative—analysis included multiple calculations using tools	Structural Composition into 1 Category (automatic)	Paired comparisons for semantic and structural model distance measures
DEEP	Annotated Influence Diagrams	Quantitative/Qualitative—analysis is done mostly by hand	Structural Decomposition into 3 Key Categories (automatic)	Unlimited comparisons, show details relating to concepts

Die *SMD-Technologie* erwies sich in der Vergleichsstudie neben weiteren drei Methodologien als zuverlässiges Verfahren zur Erfassung und Analyse externalisierter Modelle. Die Ergebnisse der *SMD-Technologie* fließen auf Grund der hohen Reliabilität und festgestellten Validität direkt in die Datenanalyse (vgl. Kapitel 8) mit ein.

7.3.6 Oberflächen-Struktur-Vergleich

Der *Oberflächen-Struktur-Vergleich* (SSC: Surface-Structure-Comparison) wurde als weiteres, leicht zu handhabendes Instrument entwickelt, um chronologisch geordnete MODELLE hinsichtlich augenscheinlicher Ähnlichkeiten zu diagnostizieren. Dafür werden zwei zu vergleichende MODELLE mittels Videoprojektor so genannten *Model-Ratern* für 5 Sekunden gezeigt. Die Gruppe der *Model-Rater* besteht aus Personen, die vor dem SSC eine Einführung in die Aufgaben und Vorgehensweise erhalten. Die *Model-Rater* erhalten die Aufgabe für eine ähnliche Struktur der projizierten MODELLE ein „+“ und für eine ungleiche Struktur ein „-“, in einen Auswertungsbogen einzutragen (vgl. Anhang A).

Das Instrument wurde in jeder der drei Studien angewendet. Die Reliabilität der SSC wurde mittels der Test-Retest-Zuverlässigkeit in den Studien 1 und 2 geprüft (vgl. Tabelle 7). Die Ergebnisse zeigen für beide Studien eine hohe Reliabilität.

Tabelle 7: Test-Retest-Reliabilität des Oberflächen-Struktur-Vergleichs

	Studie 1	Studie 2
N Model-Rater	31	21
n Modelle	5642	3822
Cronbach's Alpha	.84	.86

Mittels eines stochastischen Modells werden folglich Übergangswahrscheinlichkeiten zwischen MODELLEN bzw. zwischen jeweils zwei Messzeitpunkten berechnet (vgl. Bartholomew, 1967; Ifenthaler & Seel, 2005). Die sehr ökonomisch gewonnen Daten des *Oberflächen-Struktur-Vergleichs* enthalten wertvolle Informationen bezüglich der lernabhängigen Veränderung individueller MODELLE. Die Analyse mittels des *Oberflächen-Struktur-Vergleichs* bietet jedoch keine Informationen bezüglich der Struktur und semantischen Qualität der MODELLE.

7.4 METHODOLOGISCHE ANMERKUNGEN

Die Qualität und Aussagekraft von empirischen Forschungsstudien ist maßgeblich von der experimentellen Anordnung und den eingesetzten Diag-

noseinstrumente abhängig. Im Folgenden werden daher die Bedingungsverwirklichung der drei experimentellen Studien (Kapitel 7.4.1) und die Objektivität (Kapitel 7.4.2), Reliabilität (Kapitel 7.4.3) und Validität (Kapitel 7.4.4) der *SMD-Technologie* kritisch analysiert.

7.4.1 Kritische Analyse der Bedingungsverwirklichung

Bei der Durchführung von experimentellen und quasi-experimentellen Studien können Störfaktoren wirksam werden, welche die *interne* und *externe Validität* der Untersuchungen in Frage stellen. Die Kontrolle der *internen Validität* ermöglicht eine hinreichend eindeutige Interpretation der Versuchsergebnisse im Hinblick auf die Untersuchungshypothesen (vgl. Campbell & Stanley, 1966). In Anlehnung an die Notation von Campbell & Stanley (1966) zeigt Tabelle 8 die in den drei experimentellen Studien kontrollierten (+) bzw. unkontrollierten (-) Faktoren. Ein (?) deutet auf einen möglichen Einfluss des Faktors hin.

Tabelle 8: Überprüfung der internen Validität der experimentellen Studien nach Campbell & Stanley (1966)

History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of Selection and Maturation, etc.
?	+	+	+	+	+	-	+

Außer dem Fernbleiben von Probanden (*Mortality*) während der drei experimentellen Studien konnten alle Faktoren der internen Validität nach Campbell & Stanley (1966) kontrolliert werden. Der Faktor *History*, also die Zeit zwischen den einzelnen Messzeitpunkten des gesamten Experimentes hat eventuell einen Einfluss auf den experimentellen Ablauf, da sich die Probanden zwischen den einzelnen Messzeitpunkten treffen und Erfahrungen bzw. Ergebnisse austauschen konnten. Da dies jedoch für alle Probanden gilt, wird dieser Faktor bezüglich der internen Validität der drei experimentellen Studien als geringfügig eingestuft.

Die Kontrolle der *externen Validität* von Experimenten wird von Bracht & Glass (1975) ausführlich diskutiert und kann als Erweiterung der Darstellung von Campbell & Stanley (1966) betrachtet werden. Es wird davon ausgegangen, dass Störfaktoren der *externen Validität* die Generalisierbarkeit von experimentellen Befunden einschränken. Dabei werden zwei Arten unterschieden (Bracht & Glass, 1975, p. 65):

1. Populationsvalidität (population validity)
2. Situationsvalidität (ecological validity)

Die Stichproben der drei experimentellen Studien dieser Arbeit wurden nicht streng nach dem Zufallsprinzip aus der verfügbaren Population entnommen. Daher ist keine Repräsentativität der Stichproben anzunehmen und somit sind Verallgemeinerungen über die experimentell verfügbare Population nicht bzw. nur in sehr beschränktem Maße möglich. Die Situationsvalidität besteht dann, wenn experimentelle Effekte unabhängig von dem experimentellen Setting sind (vgl. Bracht & Glass, 1975). Die Umsetzung der drei experimentellen Studien folgte den Forderungen Browns (1992), nach denen eine experimentelle Laborsituation auch in eine *alltägliche* Lehrsituation transferierbar sein sollte. Die dreifache Replikation der experimentellen Studie, der detailliert ausgearbeitete Untersuchungsplan, das Längsschnittdesign und die exakte Einweisung der Versuchsleiter tragen in größerem Maße zur Situationsvalidität bei.

Zusammenfassend kann festgestellt werden, dass *die interne Validität* der drei experimentellen Studien als hoch eingeschätzt werden kann, da Störgrößen und weitere Effekte für alle experimentellen Variationen konstant gehalten wurden. Im Gegensatz dazu, müssen für die *externe Validität* der drei Studien erhebliche Einschränkungen festgestellt werden.

7.4.2 Objektivität

Objektivität ist dann gewährleistet, wenn die Testergebnisse vom Testanwender unabhängig sind (vgl. Bortz & Döring, 1995). Ingenkamp (1997) differenziert weiter nach einer *Durchführungs-, Auswertungs- und Interpretationsobjektivität*. Die *Durchführungsobjektivität* wird durch standardisierte Untersuchungsabläufe, Instruktionen und pädagogische Interventionen erreicht. Die Bearbeitungs-

zeiträume der multimedialen Lernumgebung und die jeweilige Konstruktion der MODELLE wurden zeitlich kontrolliert. Durch genaue Vorschriften (vgl. Kapitel 4.3) bei der Aufbereitung der MODELLE in Listenform und der automatisierten Datenanalyse kann von einer hohen *Auswertungsobjektivität* ausgegangen werden. Die *Interpretationsobjektivität* wird dadurch gewahrt, dass ein enger Bezug zur Theorie der mentalen Modelle (Seel, 1991) hergestellt wird und individuelle Auslegungen der Ergebnisse nicht akzeptiert werden (vgl. Bortz & Döring, 1995).

7.4.3 Reliabilität

Grundvoraussetzung für eine hohe *Reliabilität* ist eine hohe *Objektivität* (vgl. Bortz & Döring, 1995). In der klassischen Testtheorie wird außerdem von einer hohen *Reliabilität* ausgegangen, wenn nach wiederholter Anwendung eines Messvorgangs und bei identischen Probanden ein gemessener Wert unverändert diagnostiziert wird³³. Im Rahmen der vorliegenden empirischen Untersuchung im Längsschnitt kann eine Reliabilitätsbestimmung nach der klassischen Testtheorie nur dann erfolgen, wenn ein Lernzuwachs ausgeschlossen werden kann (siehe hierzu auch die Ausführungen in Kapitel 5). Auch Bortz & Döring (1995) weisen darauf hin, dass eine Testwiederholungsreliabilität bei zeitabhängigen Merkmalen wenig brauchbar ist.

Um dennoch eine Reliabilitätsbestimmung durchführen zu können, wurde in der dritten experimentellen Studie eine Kontrollgruppe eingeführt. Da diese Gruppe nicht in der Inhaltsdomäne der eigentlichen Problemstellung gelernt hat, kann davon ausgegangen werden, dass sich die MODELLE der Versuchspersonen der Kontrollgruppe zwischen den Messzeitpunkten nicht verändert haben.

Für die Berechnung der *Test-Retest-Reliabilität* werden die Surface-, Matching- und Deep-Structure Indizes der Messzeitpunkte drei und vier (MZP₃ und MZP₄) der Kontrollgruppe herangezogen. Als Zusammenhangsmaß wird die Rangkorrelation nach Spearman berechnet.

³³ Es gilt anzumerken, dass eine perfekte Korrelation, selbst bei zeitstabilen Persönlichkeitsmerkmalen, in den Human- und Sozialwissenschaften ausgeschlossen werden kann (vgl. Bortz & Döring, 1995).

Tabelle 9: Test-Retest-Reliabilität für die Surface-, Matching- und Deep-Structure der Kontrollgruppe (N = 10)

	Test-Retest- Reliabilität
Surface-Structure θ	.824**
Matching-Structure μ	.815**
Deep-Structure δ	.901**

** $p < .01$ (zweiseitige Signifikanz)

Die Ergebnisse in Tabelle 9 zeigen, dass ein hoch signifikanter Zusammenhang zwischen den an zwei Messzeitpunkten (MZP₃ und MZP₄) erhobenen Indizes (Surface-, Matching- und Deep-Structure) besteht. Auf der einen Seite kann dieses Ergebnis als ein deutlicher Hinweis für die Reliabilität der experimentellen Studie angesehen werden. Andererseits muss an dieser Stelle darauf hingewiesen werden, dass mentale Modelle individuelle ad hoc Konstruktionen sind (vgl. Seel, 1991) und somit eine Reliabilitätsbestimmung mittels Standardverfahren wie z. B. *Test-Retest*-, *Split-Half*- oder *Od-Even-Methode* (vgl. Rost, 2005) nur eingeschränkt erfolgen kann, da diese Verfahren von einer Stabilität der latenten Variablen ausgehen.

Der detailliert ausgearbeitete Untersuchungsplan der drei experimentellen Studien und die verwendete multimedialen Lernumgebung gewährleisten jedoch zumindest eine exakte Wiederholbarkeit der Experimente.

7.4.4 Validität

Die Bestimmung der *Objektivität* und *Reliabilität* sind keine hinreichenden Kriterien für die Gültigkeit bzw. *Validität* eines Tests bzw. Diagnoseinstruments. Trotz hoher *Reliabilität* kann ein Test durchaus unbrauchbar sein, da er ein völlig anderes Konstrukt misst, als vermutet wird (vgl. Bortz & Döring, 1995, p. 185). Rost (2005, pp. 133) führt verschiedene Arten an, wie die *Validität* bestimmt werden kann. Neben der *Inhalts*-, *Übereinstimmungs*- und *Kriteriumsvalidität* wird der *Konstruktvalidität* in Human- und Sozialwissenschaften besondere Bedeutung beigemessen (vgl. Lienert & Raatz, 1994; Dieterich, 1977, Bortz & Döring, 1995). Insbesondere bei neu konzipierten und entwickelten Diagnoseinstrumenten, wie dem der *SMD-Technologie*, ist es notwendig aus der Theorie abgeleitete Merkmale in messbare Kriterien abzubilden. Ziel der *Konstruktvalidierung* ist es, aus theoretischer Sicht zu klären, was ein Diagnosein-

strument tatsächlich misst. Hierfür stehen unterschiedliche methodologische Vorgehensweisen³⁴ zur Verfügung (vgl. Lienert & Raatz, 1994, p. 226). Kritisch gilt anzumerken, dass es kein „exaktes Maß der Höhe der Konstruktvalidität gibt“ (Lienert & Raatz, 1994, p. 228).

Die umfassende Auseinandersetzung mit der Theorie der mentalen Modelle (vgl. Johnson-Laird, 1983; Gentner & Stevens, 1983; Seel, 1991) und die Analyse verfügbarer Diagnoseinstrumente (vgl. Kapitel 3) bilden die Basis für die theoriegeleitete Entwicklung der *SMD-Technologie* (vgl. Kapitel 4). Aus empirischer Sicht soll die *Validität* der *SMD-Technologie* durch den Einsatz der Außenkriterien (1) MITOCAR und (2) deklaratives Wissen bestimmt werden.

Das von Pirnay-Dummer (2006) entwickelte Diagnoseinstrument MITOCAR (**M**odel **I**nspection **T**race **O**f **C**oncepts **A**nd **R**elations) wurde vor dem Hintergrund der Theorie der mentalen Modelle (vgl. Seel, 1991) entwickelt, um natürlichsprachige Aussagen mittels Parsing und Korpuslinguistik bezüglich deren konzeptuellen und strukturellen Ähnlichkeiten zu diagnostizieren. Die Rohdaten der experimentellen Studie 3 ($N = 47$) wurden für die MITOCAR-Software aufbereitet. Es folgte eine Diagnose der Modelle bezüglich deren konzeptuellen und strukturellen Ähnlichkeiten mittels der MITOCAR-Software. Die Analyseergebnisse von MITOCAR wurden für die Überprüfung der Validität mit den Ergebnissen der *SMD-Technologie* korreliert³⁵.

Tabelle 10: Bivariate Korrelationen nach Pearson zwischen Analyseergebnissen von MITOCAR und der SMD-Technologie für die experimentelle Studie 3 ($N = 47$)

	MITOCAR (Konzepte und Struktur)	Surface-Structure θ	Matching-Structure μ
MITOCAR (Konzepte und Struktur)	-	.610** ¹	.527** ¹
Surface-Structure θ		-	.766** ¹
Matching-Structure μ			-

** $p < .01$, * $p < .05$ (zweiseitige Signifikanz)

¹ Korrelation nach Pearson

³⁴ Korrelation eines Tests mit mehreren Außenkriterien; mit Tests ähnlichen Validitätsanspruches; mit Tests die andere Merkmale erfassen; Analyse inter- und intraindividuelle Unterschiede in den Testresultaten; Inhaltlich-logische Analyse der Testelemente; Faktorenanalyse (vgl. Lienert & Raatz, 1994, pp. 226).

Den Ergebnissen in Tabelle 10 ist zu entnehmen, dass für das Außenkriterium MITOCAR mittlere bis starke signifikante Korrelationen³⁶ zur Surface- bzw. Matching-Structure der *SMD-Technologie* bestehen³⁷.

Nachdem eine *konvergente Validität* für die *SMD-Technologie* und MITOCAR nachgewiesen werden konnte, soll die *SMD-Technologie* bezüglich eines weiteren Außenkriteriums (deklaratives Wissen) auf *divergente Validität* überprüft werden. Dabei wird angenommen, dass keine Zusammenhänge zwischen dem Surface- und Matching-Structure Index und dem deklarativen Wissen bestehen. Zwischen dem Deep-Structure Index und dem deklarativen Wissen wird ein Zusammenhang erwartet, da diese Analyseebene der *SMD-Technologie* Aussagen über die semantisch-strukturelle Ähnlichkeit der MODELLE mit dem Referenzmodell macht und somit auch Merkmale deklarativen Wissens enthalten sind.

Tabelle 11: Bivariate Korrelationen nach Pearson bzw. Spearman zwischen der SMD-Technologie und dem deklarativen Wissen für die experimentelle Studie 3 (N = 47)

	deklaratives Wissen	Surface- Structure θ	Matching- Structure μ	Deep-Structure δ
deklaratives Wissen	-	.273 ¹	.112 ¹	.355* ²
Surface-Structure θ		-	.766** ¹	.089 ²
Matching-Structure μ			-	.166 ²
Deep-Structure δ				-

** $p < .01$, * $p < .05$ (zweiseitige Signifikanz)

¹ Korrelation nach Pearson; ² Korrelation nach Spearman

Die Ergebnisse in Tabelle 11 zeigen keine Zusammenhänge zwischen dem deklarativen Wissen und der Surface- bzw. Matching-Structure der *SMD-Technologie*. Dies entspricht den theoretischen und methodologischen Überlegungen der *SMD-Technologie* (vgl. Kapitel 4), nachdem die Ebenen Surface- und

³⁵ Für die Berechnung der bivariaten Korrelation wurde jeweils das arithmetische Mittel der sieben Messzeitpunkte der dritten experimentellen Studien verwendet.

³⁶ Nach Brosius (2002) wird die Höhe des Korrelationskoeffizienten wie folgt eingeordnet: .00 (keine Korrelation); bis .20 (sehr schwache Korrelation); bis .40 (schwache Korrelation); bis .60 (mittlere Korrelation); bis .80 (starke Korrelation); bis .99 (sehr starke Korrelation); 1.00 (perfekte Korrelation).

³⁷ Der Deep-Structure Index δ der *SMD-Technologie* bezieht sich auf die semantisch-strukturelle Ähnlichkeit zwischen einem Modelle und einem Referenzmodell. Dieser Diagnoseansatz wird von MITOCAR nicht verfolgt. Daher wird der DEEP-Structure Index δ bei den Korrelationsberechnungen nicht berücksichtigt.

Matching-Structure keinen direkten Bezug zur Inhaltsdomäne zulassen. Die schwache signifikante Korrelation zwischen dem deklarativen Wissen und dem Deep-Structure Index bestätigt die theoretischen Überlegungen zur *SMD-Technologie*. Es kann davon ausgegangen werden, dass Probanden mit hohem deklarativem Wissen in einer Inhaltsdomäne auch einen hohen Deep-Structure Index δ erreichen werden.

Die empirischen Analysen zur Überprüfung der *Validität* der *SMD-Technologie* zeigen sowohl *konvergente* als auch *divergente Validität* bezüglich der verwendeten Außenkriterien. Wie bereits in Kapitel 7.3.5 dargestellt, konnte außerdem eine *Kreuzvalidierung* mit mehreren quantitativen und qualitativen Diagnoseinstrumenten zur Erfassung von externalisierten Modellen erfolgen. Dabei erwies sich die *SMD-Technologie* als zuverlässiges Verfahren zur Erfassung und Analyse externalisierter Modelle.

7.5 AUSWERTUNGSVERFAHREN

Die Analyse der erhobenen Daten basiert ausschließlich auf quantitativen Auswertungsverfahren. Für die statistische Auswertung werden die Softwarepakete SPSS 14 (SPSS, 2006) und HLM 6 student edition (SSI, 2005) verwendet. Die deskriptiven Statistiken werden mittels Maßen der zentralen Tendenz (arithmetisches Mittel M ; Median Mdn) dargestellt. Außerdem enthalten die deskriptiven Statistiken Aussagen zur Standardabweichung SD sowie die Maximal- (Max) und Minimalwerte (Min).

Vor jeder Datenauswertung erfolgt eine Prüfung der Variablen auf Normalverteilung mittels *Kolmogorov-Smirnov-Test* bzw. *Shapiro-Wilks-Test*. Bei Verletzung der Annahme auf Normalverteilung wird in der Ergebnisdarstellung darauf hingewiesen und es werden non-parametrische Testverfahren verwendet. Bei vorliegender Normalverteilung werden parametrische Testverfahren verwendet. Für alle Analysen wird eine Irrtumswahrscheinlichkeit von $\alpha = .05$ festgelegt. Werden Gruppenunterschiede nach parametrischen Testverfahren statistisch bedeutsam, so wird nach Rost (2005) zusätzlich ein Ef-

fektstärkemaß³⁸ bezüglich der praktischen Bedeutsamkeit mit angegeben. Die Interpretation der Effektstärke ist abhängig vom Effektstärkemaß (vgl. Tabelle 12):

Tabelle 12: Effektstärkemaße und deren Interpretationsrichtwerte (Rost, 2005, p. 173)

Effektstärkemaß	Kleiner Effekt	Mittlerer Effekt	Großer Effekt
r	ab 0.10	ab 0.25	ab 0.50
η^2	ab 0.01	ab 0.06	ab 0.25
Ω^2	ab 0.01	ab 0.06	ab 0.25

Wie bereits in Kapitel 5.5 dargestellt, stellt die *Hierarchisch Lineare Modellierung* (HLM) von Bryk und Raudenbush (1987, 1992) einen weiteren aussichtsreichen Ansatz zur Modellierung individueller Veränderungen dar (siehe auch Bryk, Raudenbush & Congdon, 1994, Raudenbush & Bryk, 2002). Das mathematisch-statistische Verfahren von HLM kann als Erweiterung der multiplen Regression verstanden werden, wobei die Grundlage des Verfahrens ein Mehrebenen-Regressionsmodell darstellt. Wie der Name *Hierarchisch Lineare Modellierung* verdeutlicht, wird bei diesem Verfahren von einer hierarchischen Datenstruktur ausgegangen. Die hierarchischen Ebenen für Veränderungsmessungen werden in HLM wie folgt gebildet. Bei wiederholten Messungen ist der Prädiktor der ersten Ebene (Level 1) die Zeit bzw. der jeweilige Messzeitpunkt. Die zweite Ebene (Level 2) untersucht die Ebene der Individuen. Alle *HLM-Analysen* werden mit dem Softwarepaket HLM 6 student edition (SSI, 2005) durchgeführt. Das mehrebenenanalytische Verfahren berücksichtigt individuell verschiedene Veränderungsprozesse. Daher können nicht nur gruppenspezifische Charakteristika identifiziert werden, sondern auch individuelle Verlaufskurven diagnostiziert werden (vgl. Renkl & Gruber, 1995). Das Softwarepaket HLM benötigt für die Verarbeitung der Datenanalyse eine spezielle MDM-Datei, welche aus mehreren Datensätzen, getrennt nach Level-1 und Level-2, generiert werden muss.

Weitere aussagekräftige statistische Analysemethoden wie z. B. LISREL (Jöreskog & Sörbom, 1989; Jöreskog et al., 1999) können auf Grund deren

³⁸ Nach Rost (2005, pp. 172) geben Effektstärkemaße wie Ω^2 (Omega²), η^2 (Eta²) oder der Korrelationskoeffizient r Auskunft über die Größe eines Effekts.

Anforderungen (Stichprobengröße $N < 200$) im Rahmen der vorliegenden Untersuchung nicht verwendet werden.

8 ERGEBNISSE DER UNTERSUCHUNGEN

„Bei jeder Diagnostik geht es um *Erkenntnis*, also um die Feststellung von Sachverhalten, Tatsachen, Eigenschaften oder Merkmalen, Bedingungen und dergleichen.“ (Klauer, 1978, p. 5)

Ausgangspunkt der empirischen Analysen sind die in Kapitel 6 formulierten Fragestellungen und Hypothesen und die in Kapitel 7 diskutierten methodologischen Ausführungen. Die Ergebnisdarstellung folgt der chronologischen Reihenfolge der drei experimentellen Studien. Diese gliedert sich jeweils in eine *Vorbetrachtung der Untersuchungsstichprobe*, einen *deskriptiven* und einen *hypotheseprüfenden Teil*. Abschließend werden die Ergebnisse der drei experimentellen Studien kurz zusammengefasst.

8.1 ERGEBNISSE DER ERSTEN STUDIE

An der ersten experimentellen Studie nahmen 26 Probanden (vgl. Kapitel 7.2.1) teil. Die Inhaltsdomäne entstammt aus dem Bereich *Geologie* mit dem Thema *Ökosystem Wald*. Im Hinblick auf die theoretischen Ausführungen zur *SMD-Technologie* (vgl. Kapitel 4.3.1) wurde für die erste Studie ein *Erklärungs-* und ein *Expertenmodell* (vgl. Kapitel 2.2) bezüglich der Problemstellung unter Berücksichtigung der Systemelemente Energie- und Materialfluss (von Sengenbusch, 2002) von einem Inhaltsexperten erstellt (vgl. Anhang B). Der Untersuchungsplan der ersten Studie sah acht Messzeitpunkte vor (vgl. Kapitel 7.1.2.1).

8.1.1 Vorbetrachtung der Untersuchungsstichprobe

Die Stichprobe der ersten Studie wurde bereits in Kapitel 7.2.1 kurz beleuchtet. Im Folgenden soll geprüft werden, ob sich die Untersuchungsstichprobe hinsichtlich zu berücksichtigender Merkmale unterscheidet, da dies die Aussagekraft der Ergebnisse beeinträchtigen könnte. Die Ergebnisse des Chi-Quadrat-Tests zeigen, dass sich die Untersuchungsstichprobe innerhalb der experimentellen Variation hinsichtlich der geschlechtsspezifischen Verteilung nicht unterscheidet (vgl. Tabelle 13).

Tabelle 13: Absolute Häufigkeiten und Prozentangaben der Untersuchungsstichprobe für die experimentellen Bedingungen *self-guided learning* (N = 13) und *scaffolding-based learning* (N = 13), Ergebnisse des Chi-Quadrat-Tests (Studie 1)

	self-guided learning		scaffolding-based learning		χ^2	df	p
	N	%	N	%			
Geschlecht							
weiblich	8	61.5	9	69.2			
männlich	5	38.5	4	30.8	.170	1	n.s.

Auf Grund der explorativen Datenanalyse mittels *Shapiro-Wilk-Test* wird die Annahme der Normalverteilung für die Faktoren *Alter* ($W = .762$, $df = 26$, $p = .001$) und *deklarativer Wissenstest* ($W = .908$, $df = 26$, $p = .024$) zurückgewiesen. Die Analyse der Daten erfolgt daher mit dem non-parametrischen *Mann-Whitney-U-Test* (vgl. Tabelle 14).

Tabelle 14: Mittelwerte und Standardabweichungen der Untersuchungsstichprobe für die experimentellen Bedingungen *self-guided learning* (N = 13) und *scaffolding-based learning* (N = 13), Ergebnisse des Mann-Whitney-U-Tests (Studie 1)

	self-guided learning		scaffolding-based learning		U	p
	M	SD	M	SD		
Alter	14.92	.494	15.15	.689	68.000	.316
deklaratives Wissen	13.69	2.626	13.62	2.434	83.500	.959

Die Ergebnisse des *Mann-Whitney-U-Tests* zeigen, dass sich die Untersuchungsgruppen hinsichtlich der Faktoren *Alter* und *deklaratives Wissen* (Vorwissen im Bereich Geologie) nicht signifikant unterscheiden. Weitere Datenanalysen und Interpretationen hinsichtlich der Untersuchungsgruppen sind somit unbedenklich.

8.1.2 Deskription der Ergebnisse

Die Deskription der Ergebnisse folgt den in Kapitel 6 formulierten Fragestellungen. Es werden sowohl die Ergebnisse des *Oberflächen-Struktur-Vergleiches* als auch der *SMD-Technologie* dargestellt.

8.1.2.1 Diagnose lernabhängiger Veränderung

Mittels dem *Oberflächen-Struktur-Vergleich* (vgl. Kapitel 3.2.5 und 7.3.6) wurden die MODELLE der Probanden in zeitlich chronologischer Reihenfolge von $N = 31$ „Model-Ratern“ bezüglich deren Ähnlichkeit begutachtet.

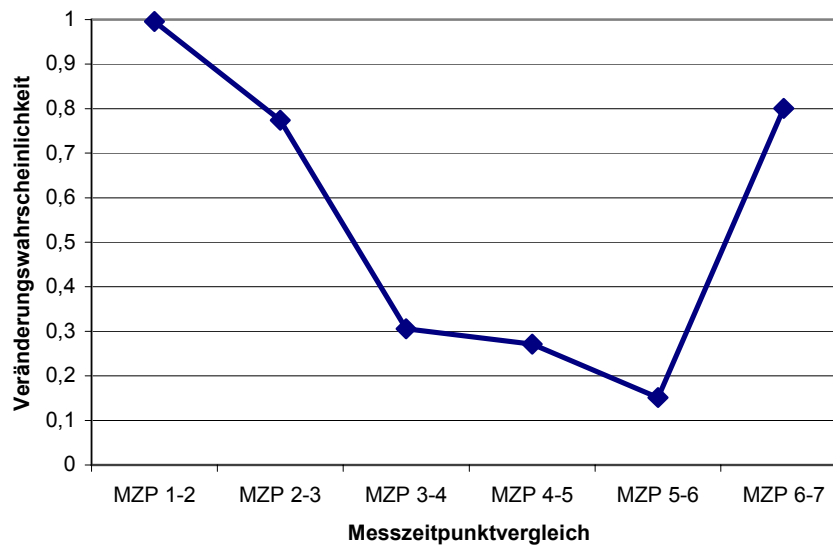


Abbildung 23: Oberflächen-Struktur-Vergleich der ersten experimentellen Studie

Wie Abbildung 23 und Tabelle 15 zu entnehmen ist, konnte mittels *Oberflächen-Struktur-Vergleich* eine hohe durchschnittliche Veränderungswahrscheinlichkeit der MODELLE zwischen dem ersten und zweiten Messzeitpunkt (MZP 1-2; 99 %, $SD = .015$) und dem zweiten und dritten Messzeitpunkt (MZP 2-3; 77 %, $SD = .337$) festgestellt werden. Ab dem dritten Messzeitpunkt ist eine geringere Veränderung der MODELLE festzustellen. Zwischen den Messzeitpunkten fünf und sechs konnte lediglich eine durchschnittliche Veränderungswahrscheinlichkeit von 15 % ($SD = .129$) festgestellt werden. Die hohe Veränderungswahrscheinlichkeit zwischen den Messzeitpunkten sechs und sieben (MZP 6-7) mit 80 % ($SD = .334$) ist durch die pädagogische Intervention *Präsentation des Lösungsmodells*³⁹ zu erklären.

³⁹ In allen drei experimentellen Studien erhielten die Probanden die pädagogische Intervention *Präsentation des Lösungsmodells*. Nach dem sechsten Messzeitpunkt MZP 6) präsentierte jeder

Tabelle 15: Durchschnittliche Veränderungswahrscheinlichkeit der MODELLE (N = 26) nach dem Oberflächen-Struktur-Vergleich (Studie 1)

Messzeitpunkt	Veränderungswahrscheinlichkeit	
	M	SD
MZP 1-2	.99547511	.0159851
MZP 2-3	.77375566	.3371742
MZP 3-4	.30542986	.2625503
MZP 4-5	.27149321	.3004389
MZP 5-6	.15158371	.1290616
MZP 6-7	.80090498	.3344476

Im Gegensatz zum *Oberflächen-Struktur-Vergleich* ermöglicht die *SMD-Technologie* (vgl. Kapitel 4) eine weitaus detailliertere Analyse der MODELLE zu den einzelnen Messzeitpunkten. Die Ergebnisse in Tabelle 16 zeigen den Surface-, Matching- und Deep-Structure Index für die einzelnen Messzeitpunkte.

Tabelle 16: Ergebnisse der SMD-Technologie für die Messzeitpunkte der ersten experimentellen Studie (N = 26)

Messzeitpunkt	Surface-Structure		Matching-Structure		Deep-Structure Erklärungsmodell		Deep-Structure Expertenmodell	
	M	SD	M	SD	M	SD	M	SD
MZP 1	5,3846	3,3356	2,0385	1,1129	0,0525	0,0702	0,0178	0,0272
MZP 2	8,3846	8,0849	1,6538	1,6957	0,0426	0,0681	0,0164	0,0280
MZP 3	19,6923	12,1351	3,0385	1,3706	0,0969	0,0898	0,0447	0,0435
MZP 4	20,9231	12,5791	3,1154	1,3950	0,1064	0,1039	0,0521	0,0585
MZP 5	21,7308	12,9478	3,2308	1,3359	0,1127	0,0928	0,0509	0,0553
MZP 6	21,4615	12,6340	3,1538	1,3767	0,1142	0,0955	0,0504	0,0556
MZP 7	11,5385	5,7427	2,8462	0,7317	0,1207	0,1016	0,0369	0,0422

Die Summe der verwendeten Propositionen (Surface-Structure) steigt zwischen dem ersten (MZP 1; $\theta = 5.39$, $SD = 3.34$) und dritten Messzeitpunkt (MZP 3; $\theta = 19.69$, $SD = 12.14$) stark an. Ab dem dritten Messzeitpunkt (MZP 3) bleibt der Surface-Structure Index nahezu konstant. Die Abnahme der verwendeten Propositionen zum letzten Messzeitpunkt (MZP 7 $\theta = 11.54$, $SD = 5.74$) kann wiederum auf die pädagogische Intervention (*Präsentation des Lösungsmodells*) zurückgeführt werden. Die lernabhängige Veränderung der Mo-

Proband im Kreis von mindestens fünf weiteren Probanden die individuelle Lösung der Problemstellung. Nach der Präsentation der individuellen Lösungen erstellten die Probanden ein letztes individuelles MODELL (MZP 7).

delle, gemessen am Surface-Structure Index der *SMD-Technologie* wird in Abbildung 24 veranschaulicht.

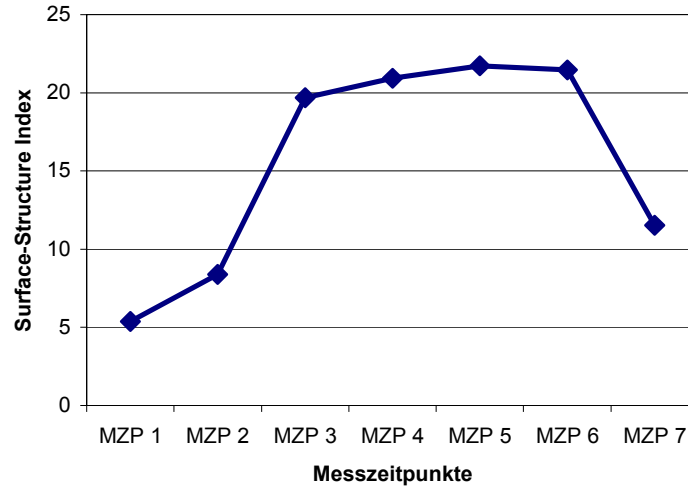


Abbildung 24: Surface-Structure Index der ersten experimentellen Studie

Bezüglich des Durchmessers bzw. der Komplexität der MODELLE (Matching-Structure) ist eine Abnahme zwischen dem ersten und zweiten Messzeitpunkt festzustellen (Tabelle 16). Zum dritten Messzeitpunkt (MZIP 3; $\mu = 3.04$, $SD = 1.37$) steigt der Matching-Structure Index an und bleibt bis zum sechsten Messzeitpunkt (MZIP 6; $\mu = 3.15$, $SD = 1.38$) nahezu konstant. Der durchschnittliche Matching-Structure Index nimmt zum siebten Messzeitpunkt (MZIP 7; $\mu = 2.85$, $SD = 0.73$) erneut ab. Der maximale Durchmesser eines MODELLS beträgt $Max = 6$. Abbildung 25 veranschaulicht die lernabhängige Veränderung der MODELLE gemessen am Matching-Structure Index der *SMD-Technologie*.

Die semantisch-strukturelle Ähnlichkeit der MODELLE zu einem Referenzmodell (Deep-Structure) wurde bezüglich zweier Referenzmodelle, einem *Erklärungs-* und einem *Expertenmodell* (vgl. Kapitel 2.2), bestimmt. Wie die Ergebnisse in Tabelle 16 zeigen, weisen die MODELLE zum ersten Messzeitpunkt (MZIP 1) nur eine geringe semantisch-strukturelle Ähnlichkeit zum *Erklärungs-* ($\delta = .05$, $SD = .07$) bzw. *Expertenmodell* ($\delta = .02$, $SD = .02$) auf. Während der Lernphase kann eine Zunahme der semantisch-strukturellen Ähnlichkeiten bezogen auf das *Erklärungsmodell* festgestellt werden.

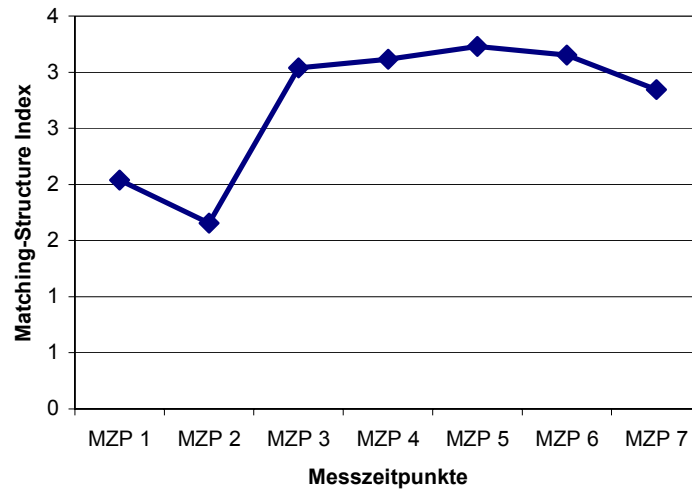


Abbildung 25: Matching-Structure Index der ersten experimentellen Studie

Der höchste Ähnlichkeitswert wird am siebten Messzeitpunkt (MZIP 7; $\delta = .12$, $SD = .10$) zwischen den MODELLEN und dem *Erklärungsmodell* erreicht. Die lernabhängige Veränderung der MODELLE, diagnostiziert an der semantisch-strukturellen Ähnlichkeit zum *Erklärungs-* bzw. *Expertenmodell* wird in Abbildung 26 veranschaulicht.

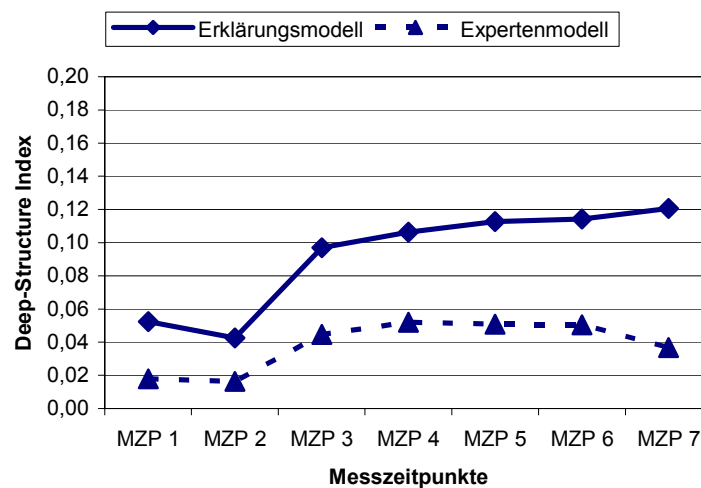


Abbildung 26: Deep-Structure Index (Erklärungs- und Expertenmodell) der ersten experimentellen Studie

8.1.2.2 Pädagogische Interventionen

Nach der allgemeinen Deskription der lernabhängigen Veränderung der Modelle werden im Folgenden die Untersuchungsgruppen (*scaffolding-based learning* und *self-guided learning*) separat betrachtet.

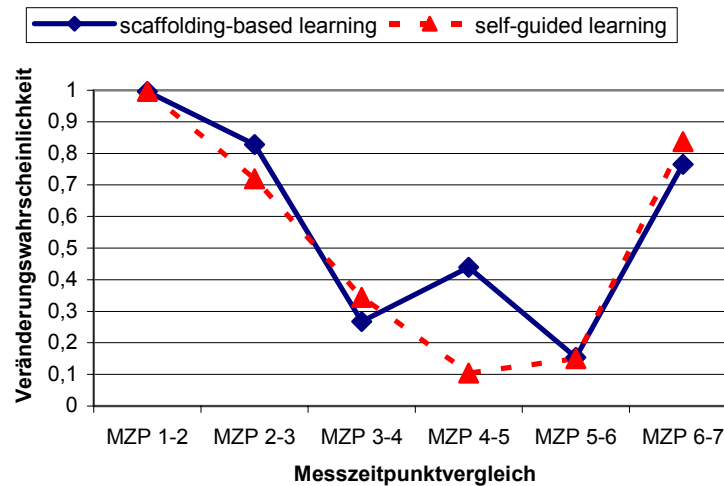


Abbildung 27: Oberflächen-Struktur-Vergleich der ersten experimentellen Studie getrennt nach den Untersuchungsgruppen

Wie Abbildung 27 zu entnehmen ist, wird mittels *Oberflächen-Struktur-Vergleich* zwischen dem ersten und zweiten Messzeitpunkt (MZIP 1-2) eine hohe durchschnittliche Veränderungswahrscheinlichkeit diagnostiziert. Die experimentelle Variation der Untersuchungsgruppen sah zwischen dem vierten und fünften Messzeitpunkt (MZIP 4-5) ein *detailliertes Feedback* (prozessorientierte Lernhilfe) für die Untersuchungsgruppe *scaffolding-based learning* vor. Im Gegensatz dazu, erhielt die Untersuchungsgruppe *self-guided learning* ein *allgemeines Feedback*⁴⁰ (vgl. Kapitel 7.1.2.1). Die Ergebnisse in Tabelle 17 zeigen, dass die Probanden der Untersuchungsgruppe *scaffolding-based learning* (43 %, $SD = .35$) ihre MODELLE zwischen dem vierten und fünften Messzeitpunkt (MZIP 4-5) deutlich mehr verändern als die Untersuchungsgruppe *self-guided learning* (10 %, SD

⁴⁰ Im Gegensatz zum *detaillierten Feedback* erhielten die Probanden mit *allgemeinem Feedback* lediglich Hinweise zum Vorgehen in der multimedialen Lernumgebung. Das *detaillierte Feedback* bezog sich auf die zum vierten Messzeitpunkt (MZIP 4) konstruierten MODELLE.

= .06). Da alle Probanden zwischen dem sechsten und siebten Messzeitpunkt die pädagogische Intervention *Präsentation des Lösungsmodells* erhielten, sind die Veränderungen zwischen den Untersuchungsgruppen nahezu identisch.

Tabelle 17: Durchschnittliche Veränderungswahrscheinlichkeit der MODELLE getrennt nach den Untersuchungsgruppen (Oberflächen-Struktur-Vergleich; Studie 1)

Messzeitpunkt	scaffolding-based learning		self-guided learning	
	M	SD	M	SD
MZP 1-2	.9954751	.0163147	.9954751	.0163147
MZP 2-3	.8280543	.2955471	.719457	.3782997
MZP 3-4	.2669683	.2615447	.3438914	.2683262
MZP 4-5	.438914	.3509876	.1040724	.0642311
MZP 5-6	.1538462	.1348639	.1493213	.1284622
MZP 6-7	.7647059	.3849834	.8371041	.286322

Eine detaillierte Ergebnisdarstellung der mittels *SMD-Technologie* ermittelten Daten erfolgt getrennt nach den Untersuchungsgruppen *scaffolding-based learning* und *self-guided learning* für die Surface-, Matching- und Deep-Structure im Anhang F.

Die durchschnittliche Anzahl der verwendeten Propositionen (Surface-Structure) steigt bei beiden Untersuchungsgruppen nach dem ersten Messzeitpunkt (MZP 1) deutlich an (vgl. Abbildung 28). Die Probanden unter der Bedingung *self-guided learning* erreichen am fünften Messzeitpunkt den höchsten durchschnittlichen Wert (MZP 5; $\theta_{\text{self-guided learning}} = 23.69$, $SD = 16.71$). Auch die Probanden der Untersuchungsgruppe *scaffolding-based learning* erreichen am fünften Messzeitpunkt den durchschnittlich höchsten Wert (MZP 5; $\theta_{\text{scaffolding-based learning}} = 19.77$, $SD = 7.85$). Der Unterschied am fünften Messzeitpunkt (MZP 5) ist jedoch nicht signifikant ($t_{(24)} = .766$, $p = .451$). Die Untersuchungsgruppe *scaffolding-based learning* bleibt jedoch insgesamt unter den Werten der Probanden mit der experimentellen Bedingung *self-guided learning*. Zwischen den Messzeitpunkten vier (MZP 4; $\theta_{\text{scaffolding-based learning}} = 18.15$, $SD = 5.81$) und fünf (MZP 5; $\theta_{\text{scaffolding-based learning}} = 19.77$, $SD = 7.85$) gilt es einen Anstieg der durchschnittlich verwendeten Propositionen der Untersuchungsgruppe *scaffolding-based learning* zu berichten. Die Surface-Structure der Untersuchungsgruppe *self-guided learning* bleibt zwischen diesen Untersuchungszeitpunkten konstant (MZP 4 & 5; $\theta_{\text{self-guided learning}} = 23.69$, $SD = 16.71$).

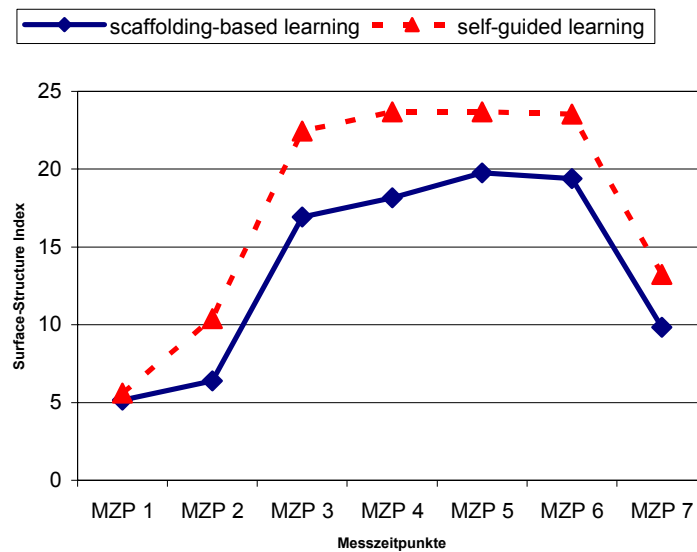


Abbildung 28: Surface-Structure Index der ersten experimentellen Studie getrennt nach den Untersuchungsgruppen

Hinsichtlich der Komplexität der MODELLE (Matching-Structure) zeigen die Ergebnisse, dass die MODELLE der Untersuchungsgruppe *self-guided learning* einen durchschnittlich größeren Durchmesser vorweisen, als die der Untersuchungsgruppe *scaffolding-based learning* (vgl. Anhang F und Abbildung 29). Mit-

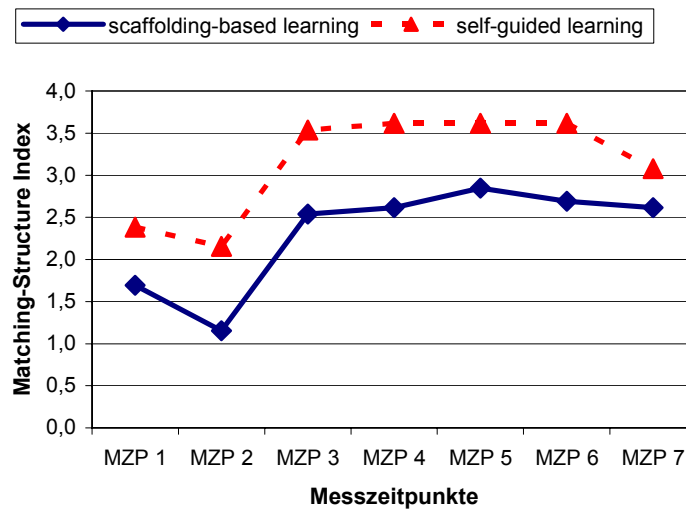


Abbildung 29: Matching-Structure Index der ersten experimentellen Studie getrennt nach den Untersuchungsgruppen

tels einer t-Test-Analyse konnten jedoch keine signifikanten Unterschiede zwischen den Untersuchungsgruppen festgestellt werden. Zum Zeitpunkt der pädagogischen Intervention (MZIP 4 - 5) vergrößert sich der Durchmesser der Modelle der Untersuchungsgruppe *scaffolding-based learning*. Der Durchmesser der Untersuchungsgruppe *self-guided learning* bleibt zu diesem Zeitpunkt konstant (vgl. Anhang F).

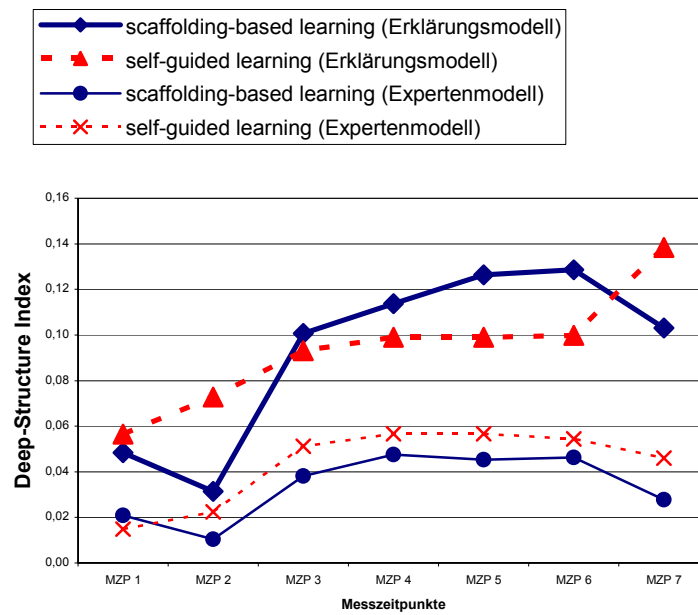


Abbildung 30: Deep-Structure Index der ersten experimentellen Studie getrennt nach den Untersuchungsgruppen und Ähnlichkeit zum Erklärungs- bzw. Expertenmodell

Die Veränderung der semantisch-strukturellen Ähnlichkeit der MODELLE zum *Erklärungs-* bzw. *Expertenmodell* (Deep-Structure) wird in Abbildung 30 veranschaulicht. Die MODELLE der Untersuchungsgruppe *scaffolding-based learning* weisen zum zweiten Messzeitpunkt sowohl für das *Erklärungs-* (MZIP 2; $\delta_{\text{scaffolding-based learning}} = .03, SD = .06$) als auch für das *Expertenmodell* (MZIP 2; $\delta_{\text{scaffolding-based learning}} = .01, SD = .02$) die niedrigste semantisch-strukturelle Ähnlichkeit während der Lernphase auf. Für die Untersuchungsgruppe *self-guided learning* konnte die größte Ähnlichkeit zum *Erklärungsmodell* (MZIP 7; $\delta_{\text{self-guided learning}} = .14, SD = .10$) am siebten Messzeitpunkt diagnostiziert werden. Bezogen auf das *Erklärungsmodell* erreicht die Untersuchungsgruppe *scaffolding-based learning* den größten Deep-Structure Index (MZIP 6; $\delta_{\text{scaffolding-based learning}} = .13, SD = .11$)

am sechsten Messzeitpunkt. Die Ergebnisse der einzelnen Messzeitpunkte werden getrennt nach den Untersuchungsgruppen in Anhang F tabellarisch dargestellt. Analysen mittels t-Test zeigen, dass sich die Untersuchungsgruppen zu keinem Messzeitpunkt signifikant unterscheiden.

8.1.2.3 Effekte von Lernervoraussetzungen

Die Lernervoraussetzungen *deklaratives Wissen*, *Lernstrategien* und *allgemeine kognitive Fähigkeiten* wurden mittels der in Kapitel 7.3 dargestellten Instrumente zu Beginn der experimentellen Studie erhoben. Tabelle 18 veranschaulicht die Ergebnisse getrennt nach den Untersuchungsgruppen.

Tabelle 18: Skalenwerte für deklaratives Wissen, Lernstrategien und allgemeine kognitive Fähigkeiten getrennt nach den Untersuchungsgruppen (Studie 1)

	self-guided learning		scaffolding-based learning	
	M	SD	M	SD
deklaratives Wissen	13.69	2.626	13.62	2.434
<i>Lernstrategien</i>				
Organisieren (LIST)	2.93	.698	3.11	.589
Elaborieren (LIST)	3.53	.468	3.45	.546
Kritisches Prüfen (LIST)	3.10	.499	3.25	.630
Wiederholen (LIST)	3.25	.374	3.21	.637
Metakognitive Strategien (LIST)	3.52	.495	3.55	.397
<i>allgemeine kognitive Fähigkeiten</i>				
Einfallsreichtum (BIS)	90.077	11.199	90.615	8.231
sprachgebundenes Denken (BIS)	97.000	11.626	94.539	14.246
figural-bildhaftes Denken (BIS)	95.077	12.332	94.308	9.835

Die Untersuchungsgruppe *scaffolding-based learning* löste von den insgesamt 18 Multiple-Choice-Fragen durchschnittlich $M = 13.62$ ($SD = 2.434$) Fragen richtig. Ähnlich viele Antworten wurden durch die Untersuchungsgruppe *self-guided learning* richtig gelöst ($M = 13.69$, $SD = 2.626$).

Die *Lernstrategien* der Probanden wurden mittels LIST-Test (vgl. Kapitel 7.3) erhoben. Die Ausprägungen der Skalen werden durch eine 5-stufigen Ratingskala (1 = sehr selten, 2 = selten, 3 = manchmal, 4 = oft, 5 = sehr oft) gemessen. Die Probanden der Untersuchungsgruppe *scaffolding-based learning*

gaben für die Skala *Metakognitive Strategien* den höchsten durchschnittlichen Wert an ($M = 3.55$, $SD = .397$). Der niedrigste durchschnittliche Skalenwert wurde von der Untersuchungsgruppe *self-guided learning* für *Organisieren* ($M = 2.93$, $SD = .698$) angegeben.

Die *allgemeinen kognitiven* Fähigkeiten der Probanden wurden durch relevante Skalen des BIS-Tests (vgl. Kapitel 7.3) diagnostiziert. Den höchsten durchschnittlichen Wert erreicht die Untersuchungsgruppe *self-guided learning* für die Skala *sprachgebundenes Denken* ($M = 97.00$, $SD = 11.626$). Niedrigster durchschnittlicher Wert wird durch die Untersuchungsgruppe *scaffolding-based learning* in der Skala *Einfallsreichtum* ($M = 90.615$, $SD = 8.231$) erreicht.

8.1.3 Hypothesenprüfende Darstellung

Die mittels Gruppenmittelwertsdiagrammen dargestellten, deskriptiven Ergebnisse in Kapitel 8.1.2 sind auf Grund ihres eingeschränkten Informationsgehaltes für die Fragestellungen bezüglich der lernabhängigen Veränderung der MODELLE nicht ausreichend. Die weiteren Analysen erfolgen daher mittels HLM (Bryk & Raudenbusch, 1987, 1992; Bryk et al., 1994; Raudenbush & Bryk, 2002). Das mehrebenenanalytische Regressionsmodell von HLM ermöglicht die Analyse von individuellen Veränderungen über mindestens drei Messzeitpunkte⁴¹. Die hypothesenprüfende HLM-Analyse erfolgt in zwei Auswertungsschritten. Zunächst wird in einem ersten Wachstumsmodell (*Level-1* Modell; Gleichung 8.1.3.1) die intraindividuelle Veränderung bezüglich der abhängigen Variablen (Surface-, Matching- und Deep-Structure Indizes der *SMD-Technologie*) geprüft.

$$Y_{ti} = \pi_{0i} + \pi_{1i}(TIME) + e_{ti} \quad [8.1.3.1]$$

Dabei ist π_{1i} der Wachstumsparameter der Person i , dargestellt als linearer Zeiteffekt ($TIME$).

⁴¹ Da für alle drei experimentellen Studien sieben Messzeitpunkte vorgesehen waren, kann eine HLM-Analyse ohne Einschränkungen erfolgen.

Die Ergebnisse in Tabelle 19 zeigen, dass die durchschnittliche Merkmalsausprägung des Surface-Structure Index zum ersten Messzeitpunkt (MZIP 1) mit 10.58 signifikant über 0 liegt⁴².

Die signifikant positiv lineare Veränderung entspricht einem Zuwachs des Surface-Structure Index von 1.66 je Messzeitpunkt. Das *Level-1* Modell klärt 50,7 % der Varianz auf. Somit kann $H_{1[1.1]}$ angenommen werden, welche besagt, dass sich die Anzahl der Propositionen (Surface-Structure) der MODELLE während des Lernprozesses signifikant verändert.

Tabelle 19: Lineares Wachstumsmodell der Surface-Structure (Studie 1)

Fixe Effekte	<i>Coefficient</i>	<i>SE</i>	<i>t-Wert</i>	<i>df</i>	<i>p</i>
Mean Initial Status π_{0i}	10.589	1.540	6.872	25	< .001
Mean Growth Rate π_{1i}	1.666	0.378	4.402	25	< .001
Zufällige Effekte	<i>Variance Component</i>	<i>SD</i>	<i>df</i>	χ^2	<i>p</i>
Initial Status r_{0i}	19.617	4.429	25	20.366	> .500
Growth Rate r_{1i}	0.484	0.695	25	10.695	> .500
Level-1 Error e_{ij}	90.718	9.524			

Ferner zeigt Tabelle 19, dass sich keine signifikanten Effekte bezüglich der interindividuellen Unterschiede im Ausgangsniveau und bezüglich des Wachstumsparameters ergaben. Die Probanden unterscheiden sich demnach nicht bezüglich ihrer Surface-Structure Indizes ($\chi^2 = 20.366$; $p > .500$) zu Beginn der ersten experimentellen Studie. Ebenso wenig unterscheiden sich die Probanden hinsichtlich der Veränderung des Surface-Structure Index während der Lernphase der ersten experimentellen Studie ($\chi^2 = 10.695$; $p > .500$).

⁴² Renkl & Gruber (1995, p. 184) bezeichnen dies als „trivialer Effekt“. Dennoch soll auf die Angabe dieser Effekte nicht verzichtet werden.

Tabelle 20: Lineares Wachstumsmodell der Matching-Structure (Studie 1)

Fixe Effekte	<i>Coefficient</i>	<i>SE</i>	<i>t-Wert</i>	<i>df</i>	<i>p</i>
Mean Initial Status π_{0i}	2.123	0.201	10.540	25	< .001
Mean Growth Rate π_{1i}	0.200	0.040	4.970	25	< .001

Zufällige Effekte	<i>Variance Component</i>	<i>SD</i>	<i>df</i>	χ^2	<i>p</i>
Initial Status r_{0i}	0.532	0.729	25	44.685	.009
Growth Rate r_{1i}	0.002	0.045	25	19.025	> .500
Level-1 Error e_{ti}	1.126	1.061			

Die durchschnittliche Merkmalsausprägung des Matching-Structure Index zum ersten Messzeitpunkt (MZIP 1) liegt mit 2.12 signifikant über 0. Pro Messzeitpunkt kann eine signifikant positive Veränderung des Matching-Structure Index von 0.20 festgestellt (vgl. Tabelle 20). Das *Level-1* Modell klärt 26,0 % der Varianz auf. Folglich kann $H_{1[1,2]}$ angenommen werden; der Durchmesser (Matching-Structure) der Modelle verändert sich während des Lernprozesses. Auf interindividueller Ebene kann ein signifikanter Unterschied zum ersten Messzeitpunkt (MZIP 1) festgestellt werden ($\chi^2=44.685$; $p=.009$). Die Parameter der Wachstumskurve weisen jedoch keine signifikanten interindividuellen Unterschiede auf ($\chi^2=19.025$; $p>.500$).

Tabelle 21: Lineares Wachstumsmodell der Deep-Structure, Erklärungsmodell (Studie 1)

Fixe Effekte	<i>Coefficient</i>	<i>SE</i>	<i>t-Wert</i>	<i>df</i>	<i>p</i>
Mean Initial Status π_{0i}	0.053	0.010	5.123	25	< .001
Mean Growth Rate π_{1i}	0.012	0.003	4.290	25	< .001

Zufällige Effekte	<i>Variance Component</i>	<i>SD</i>	<i>df</i>	χ^2	<i>p</i>
Initial Status r_{0i}	0.00107	0.03272	25	38.936	.037
Growth Rate r_{1i}	0.00010	0.01018	25	42.941	.014
Level-1 Error e_{ti}	0.00376	0.06136			

Tabelle 21 zeigt, dass zum ersten Messzeitpunkt (MZIP 1) im Gruppendurchschnitt die Ähnlichkeit zum Erklärungsmodell von 0.05 signifikant über 0 liegt. Außerdem ist die positiv, lineare Veränderung pro Messzeitpunkt um 0.01 signifikant. Das *Level-1* Modell klärt 71,8 % der Varianz auf. Die Hypothe-

se $H_{1[1.3]}$ wird angenommen; die Ähnlichkeit der MODELLE zum Erklärungsmodell (Deep-Structure) verändert sich während des Lernprozesses. In Bezug auf die interindividuellen Unterschiede im Ausgangsniveau können signifikante Effekte festgestellt werden ($\chi^2 = 38.936$; $p = .037$). Für die Parameter der Wachstumskurve ergaben sich ebenfalls signifikante interindividuelle Unterschiede ($\chi^2 = 42.941$; $p = .014$). Die Probanden unterscheiden sich demnach signifikant hinsichtlich ihres Ausgangsniveaus und hinsichtlich ihrer lernabhängigen Veränderung des Deep-Structure Index (Ähnlichkeit zum Erklärungsmodell).

Die durchschnittliche Merkmalsausprägung des Deep-Structure Index (Expertenmodell) zum ersten Messzeitpunkt (MZP 1) liegt mit 0.02 signifikant über 0. Pro Messzeitpunkt wurde eine signifikant positive Veränderung des Deep-Structure Index (Expertenmodell) von 0.004 festgestellt (vgl. Tabelle 22; 81,9 % aufgeklärte Varianz). Folglich wird Hypothese $H_{1[1.3]}$ angenommen; die Ähnlichkeit der MODELLE mit dem Expertenmodell (Deep-Structure) verändert sich während des Lernprozesses. Auf interindividueller Ebene kann kein signifikanter Unterschied zum ersten Messzeitpunkt (MZP 1) festgestellt werden ($\chi^2 = 32.079$; $p = .156$). Die Parameter der Wachstumskurve weisen jedoch signifikante interindividuelle Unterschiede auf ($\chi^2 = 58.707$; $p < .001$). Probanden unterscheiden sich somit auch signifikant hinsichtlich ihrer positiven Veränderung des Ähnlichkeitsindex zum Expertenmodell (Deep-Structure, Expertenmodell).

Tabelle 22: Lineares Wachstumsmodell der Deep-Structure, Expertenmodell (Studie 1)

Fixe Effekte	<i>Coefficient</i>	<i>SE</i>	<i>t-Wert</i>	<i>df</i>	<i>p</i>
Mean Initial Status π_{0i}	0.024	0.004	5.143	25	< .001
Mean Growth Rate π_{1i}	0.004	0.001	2.753	25	.011

Zufällige Effekte	<i>Variance Component</i>	<i>SD</i>	<i>df</i>	χ^2	<i>p</i>
Initial Status r_{0i}	0.00021	0.01438	25	32.079	.156
Growth Rate r_{1i}	0.00005	0.00680	25	58.707	< .001
Level-1 Error e_{ij}	0.00081	0.02855			

Mit dem zweiten Wachstumsmodell (*Level-2 Modell*) sollen mögliche Einflüsse von zeitstabilen Drittvariablen aufgeklärt werden (vgl. Raudenbush &

Bryk, 2002). Hierfür wird das *Level-2* Submodell nach der allgemeinen Gleichung 8.1.3.2 konstruiert:

$$\begin{aligned}\pi_{0i} &= \gamma_{00} + \gamma_{01}PRÄDIKTOR + \xi_{0i} \\ \pi_{1i} &= \gamma_{10} + \gamma_{11}PRÄDIKTOR + \xi_{1i}\end{aligned}\quad [8.1.3.2]$$

Die Modellgleichung berücksichtigt neben den Residuen ξ_{0i} und ξ_{1i} , den Intercept π_{0i} und den Steigungskoeffizienten π_{1i} , der individuellen Veränderungsverläufe, welche durch den *PRÄDIKTOR* erklärt werden.

Für jede der Ebenen der *SMD-Technologie* (Surface-, Matching- und Deep-Structure) wurde eine HLM-Analyse nach Gleichung 8.1.3.2 mit den Prädiktoren *GRUPPE* (0 = *self-guided learning*; 1 = *scaffolding-based learning*) und *VORWISSEN*⁴³ (0 = *niedrigeres Vorwissen*; 1 = *höheres Vorwissen*) durchgeführt. Die Ergebnisse werden in Tabellenform in Anhang F dargestellt. Hinsichtlich der Surface-Structure und der experimentellen Variation *GRUPPE* ergibt sich mittels Substitution der Ergebnisse folgendes *Level-2* Submodell (Gleichung 8.1.3.3):

$$\begin{aligned}\hat{\pi}_{0i} &= 15.175 - 3.057GRUPPE_i \\ \hat{\pi}_{1i} &= 2.065 - 0.266GRUPPE_i\end{aligned}\quad [8.1.3.3]$$

Aus dem ersten Teil des *Level-2* Submodells (Gleichung 8.1.3.3) ist zu entnehmen, dass der wahre Wert der Surface-Structure zum ersten Messzeitpunkt (MZP 1) im Gruppendurchschnitt der Untersuchungsgruppe *self-guided learning* 15.175 beträgt; für die Untersuchungsgruppe *scaffolding-based learning* wird dieser Wert um -3.057 niedriger geschätzt (12.118). Die lernabhängige Veränderung des Surface-Struture Index wird im zweiten Teil der Gleichung 8.1.3.3 angegeben. Demnach verändert sich der Surface-Structure Index der Probanden unter der Bedingung *self-guided learning* um 2.065; für Probanden der Bedingung *scaffolding-based learning* fällt die Veränderung des Surface-Structure Index pro Messzeitpunkt geringer aus (2.065 - 0.266 = 1.799). Wie der Ergebnisdarstellung in Anhang F zu entnehmen ist, kann nur für den ersten Messzeitpunkt ein signifikanter intraindividueller Unterschied innerhalb der Bedingung *self-guided learning*

⁴³ Die Variable *deklaratives Vorwissen* wurde durch Mediansplit (0 = niedrigeres Vorwissen; 1 = höheres Vorwissen) aufgeteilt.

festgestellt werden. Alle anderen Gruppenunterschiede sind nicht signifikant. Das *Level-2* Modell klärt 53,9 % der Varianz auf. Die Hypothese $H_{0[2.1]}$ wird beibehalten. Die lernabhängige Veränderung der Anzahl der Propositionen (Surface-Structure) der MODELLE unterscheidet sich nicht zwischen den Untersuchungsgruppen *self-guided learning* und *scaffolding-based learning*.

Die Ergebnisse bezüglich des Prädiktors *VORWISSEN* zeigen, dass sich die Probanden mit *niedrigerem Vorwissen* zum ersten Messzeitpunkt (MZIP 1) signifikant von 0 unterscheiden und sich während des Lernprozesses signifikant verbesserten. Anhang F ist weiter zu entnehmen, dass sich die Probanden mit *höherem Vorwissen* bezüglich dem ersten Messzeitpunkt (MZIP 1) und der lernabhängigen Veränderung nicht signifikant unterscheiden (48,3 % aufgeklärte Varianz). Die Hypothese $H_{0[3.1]}$ wird beibehalten, welche besagt, dass das bereichsspezifische Vorwissen keinen Einfluss auf die lernabhängige Veränderung der Anzahl der Propositionen (Surface-Structure) hat.

Die Ergebnisse der HLM-Analysen für die Matching-Structure werden in Anhang F berichtet. Aus der allgemeinen Formel des *Level-2* Submodells ergibt sich durch Substitution folgende Gleichung (8.1.3.4):

$$\begin{aligned}\hat{\pi}_{0i} &= 3.549 - 0.950GRUPPE_i \\ \hat{\pi}_{1i} &= 0.142 + 0.038GRUPPE_i\end{aligned}\quad [8.1.3.4]$$

Hinsichtlich des Prädiktors *GRUPPE* gilt es signifikante Unterschiede für den ersten Messzeitpunkt (MZIP 1) für die Untersuchungsgruppen (*self-guided learning* und *scaffolding-based learning*) zu berichten. Die lernabhängige Veränderung der beiden Untersuchungsgruppen ist jedoch nicht signifikant (52% aufgeklärte Varianz). Folglich wird die Hypothese $H_{0[2.2]}$ beibehalten; die lernabhängige Veränderung des Durchmesser (Matching-Structure) der MODELLE unterscheidet sich nicht zwischen den Untersuchungsgruppen.

Für den Prädiktor *VORWISSEN* konnte ein signifikanter Effekt für die Untersuchungsgruppe mit *niedrigerem Vorwissen* zum ersten Messzeitpunkt (MZIP 1) und hinsichtlich deren lernabhängigen Veränderung festgestellt werden (vgl. Anhang F; 23,1 % aufgeklärte Varianz). Die Probanden mit *höherem Vorwissen* unterschieden sich jedoch nicht bezüglich dem Wert zum ersten Messzeitpunkt (MZIP 1) und deren lernabhängigen Veränderung der Matching-Structure. Somit wird die Hypothese $H_{0[3.2]}$ beibehalten. Demnach hat das be-

reichsspezifische Vorwissen keinen Einfluss auf die lernabhängige Veränderung des Durchmesser (Matching-Structure).

In einer weiteren HLM-Analyse bezüglich der Deep-Structure (Erklärungsmodell) wurde folgendes *Level-2* Submodell ermittelt (Gleichung 8.1.3.5):

$$\begin{aligned}\hat{\pi}_{0i} &= 0.057 - 0.002GRUPPE_i \\ \hat{\pi}_{1i} &= 0.010 + 0.001GRUPPE_i\end{aligned}\quad [8.1.3.5]$$

Die Ergebnisse in Anhang F zeigen keine signifikanten Unterschiede für die Untersuchungsgruppen (*self-guided learning* und *scaffolding-based learning*) zum ersten Messzeitpunkt (MZIP 1) und bezüglich der lernabhängigen Veränderung auf (70,2 % aufgeklärte Varianz). Die Hypothese $H_{0[2,3]}$ wird beibehalten; die lernabhängige Veränderung der Ähnlichkeit der MODELLE mit dem Erklärungsmodell (Deep-Structure) unterscheidet sich nicht zwischen den Untersuchungsgruppen *self-guided learning* und *scaffolding-based learning*.

Der Prädiktor *VORWISSEN* zeigt, dass sich die Probanden mit *niedrigerem* und *höherem* Vorwissen sowohl zum ersten Messzeitpunkt (MZIP 1) als auch bezüglich der lernabhängigen Veränderung nicht signifikant unterscheiden (vgl. Anhang F; 72 % aufgeklärte Varianz). Die Hypothese $H_{0[3,3]}$ wird beibehalten, welche besagt, dass das bereichsspezifische Vorwissen in keinem Zusammenhang mit der lernabhängigen Veränderung der Ähnlichkeit eines konstruierten MODELLS zum Erklärungsmodell (Deep-Structure) steht.

Eine weitere HLM-Analyse (Prädiktor *GRUPPE*) ergab für die Deep-Structure (Expertenmodell) folgendes *Level-2* Submodell (Gleichung 8.1.3.6):

$$\begin{aligned}\hat{\pi}_{0i} &= 0.028 - 0.002GRUPPE_i \\ \hat{\pi}_{1i} &= 0.008 - 0.002GRUPPE_i\end{aligned}\quad [8.1.3.6]$$

Die Ergebnisdarstellung in Anhang F zeigt keine signifikanten Unterschiede für die Untersuchungsgruppen (*self-guided learning* und *scaffolding-based learning*) zum ersten Messzeitpunkt (MZIP 1) und bezüglich der lernabhängigen Veränderung auf (81,9 % aufgeklärte Varianz). Die Hypothese $H_{0[2,3]}$ wird beibehalten; die lernabhängige Veränderung der Ähnlichkeit der MODELLE mit dem Expertenmodell (Deep-Structure) unterscheidet sich nicht zwischen den Untersuchungsgruppen.

Auch hinsichtlich des Prädiktors *VORWISSEN* können keine signifikanten Ergebnisse bezüglich der lernabhängigen Veränderung berichtet werden (vgl. Anhang F, 81,9 % aufgeklärte Varianz). Somit wird die Hypothese $H_{0[3.3]}$ beibehalten, welche besagt, dass das bereichsspezifische Vorwissen in keinem Zusammenhang mit der lernabhängigen Veränderung der Ähnlichkeit eines konstruierten MODELLS zum Expertenmodell (Deep-Structure) steht.

8.2 ERGEBNISSE DER ZWEITEN STUDIE

Für die zweite experimentelle Studie konnten 33 Probanden (vgl. Kapitel 7.2.2) gewonnen werden. Analog zur ersten Studie sah der Untersuchungsplan acht Messzeitpunkte vor (vgl. Kapitel 7.1.2.2). Als Inhaltsdomäne wurde der Bereich *Geophysik* mit dem Thema *Plattentektonik* gewählt. Das zur Problemstellung passende *Erklärungs-* und *Lösungsmodell* wurde von einem Inhaltsexperten erstellt (vgl. Anhang D).

8.2.1 Vorbetrachtung der Untersuchungsstichprobe

Die bereits in Kapitel 7.2.2 beschriebene Stichprobe der zweiten experimentellen Studie soll im Folgenden hinsichtlich zu berücksichtigender Merkmale, die die Aussagekraft der Ergebnisse beeinträchtigen könnten, geprüft werden. In Tabelle 23 zeigen die Ergebnisse des Chi-Quadrat-Tests, dass sich die Untersuchungsstichprobe innerhalb der experimentellen Variation hinsichtlich der geschlechtsspezifischen Verteilung nicht unterscheiden.

Tabelle 23: Absolute Häufigkeiten und Prozentangaben der Untersuchungsstichprobe für die experimentellen Bedingungen *self-guided learning* (N = 16) und *scaffolding-based learning* (N = 17), Ergebnisse des Chi-Quadrat-Tests (Studie 2)

	self-guided learning		scaffolding-based learning		χ^2	df	p
	N	%	N	%			
Geschlecht							
weiblich	9	56.3	7	41.2			
männlich	7	43.7	10	58.8	.750	1	n.s.

Mittels explorativer Datenanalyse durch den *Shapiro-Wilk-Test* wird für den Faktor *deklarativer Wissenstest* eine normalverteilte Datengrundlage bestätigt. Für

den Faktor *Alter* ($W = .904$, $df = 33$, $p = .007$) muss jedoch die Annahme auf Normalverteilung zurückgewiesen werden.

Tabelle 24: Mittelwerte und Standardabweichungen der Untersuchungsstichprobe für die experimentellen Bedingungen *self-guided learning* (N = 16) und *scaffolding-based learning* (N = 17), Ergebnisse des Mann-Whitney-U-Tests (Studie 2)

	self-guided learning		scaffolding-based learning		U	p
	M	SD	M	SD		
Alter	14.19	1.109	14.24	1.147	136.00	n.s.

Wie in Tabelle 24 dargestellt, unterscheiden sich die Untersuchungsgruppen hinsichtlich des Faktors *Alter* nicht signifikant. Auch der *t-Test* in Bezug auf den Faktor *deklaratives Wissen* (Vorwissen im Bereich Geographie / Geophysik) zeigt keine signifikanten Ergebnisse (vgl. Tabelle 25).

Tabelle 25: Mittelwerte und Standardabweichungen der Untersuchungsstichprobe für die experimentellen Bedingungen *self-guided learning* (N = 16) und *scaffolding-based learning* (N = 17), Ergebnisse des t-Tests (Studie 2)

	self-guided learning		scaffolding-based learning		t	df	p
	M	SD	M	SD			
deklaratives Wissen	29.799	3.936	15.158	3.242	1.119	31	.272

8.2.2 Deskription der Ergebnisse

Die Deskription der Ergebnisse der zweiten Studie folgt den formulierten Fragestellungen (Kapitel 6) sowohl für den *Oberflächen-Struktur-Vergleich* als auch der *SMD-Technologie*.

8.2.2.1 Diagnose lernabhängiger Veränderung

Ähnlich wie in der ersten experimentellen Studie konnte für die zweite experimentelle Studie mittels *Oberflächen-Struktur-Vergleich* ($N = 23$ Model-Rater) eine hohe durchschnittliche Veränderungswahrscheinlichkeit für die MODELLE zwischen dem ersten und zweiten Messzeitpunkt (MZP 1-2; 87 %; $SD = .254$) festgestellt werden (vgl. Abbildung 31).

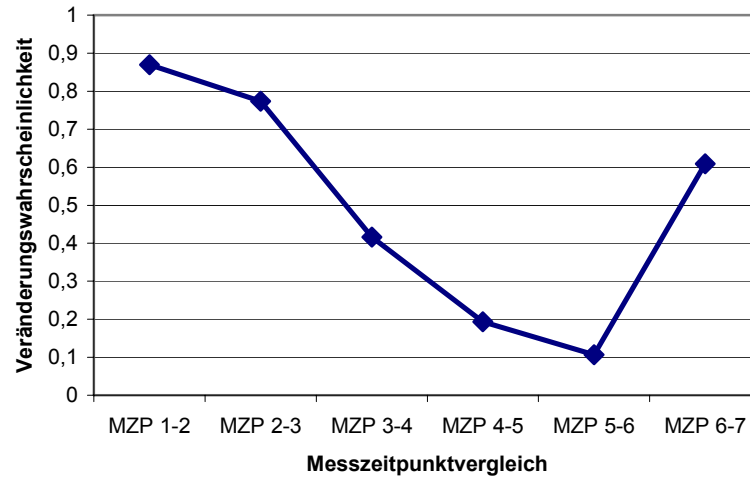


Abbildung 31: Oberflächen-Struktur-Vergleich der zweiten experimentellen Studie

Die niedrigste durchschnittliche Veränderungswahrscheinlichkeit zwischen den Messzeitpunkten fünf und sechs beträgt 11 % (MZIP 5-6; $SD = .165$). Die pädagogische Intervention *Präsentation des Lösungsmodells* erfolgte zwischen den Messzeitpunkten sechs und sieben, wodurch die erneut hohe durchschnittliche Veränderungswahrscheinlichkeit (MZIP 6-7; 61 %; $SD = .333$) zu erklären ist (vgl. Tabelle 26).

Tabelle 26: Durchschnittliche Veränderungswahrscheinlichkeit der MODELLE (N = 33) nach dem Oberflächen-Struktur-Vergleich (Studie 2)

Messzeitpunkt	Veränderungs- wahrscheinlichkeit	
	M	SD
MZIP 1-2	.86956522	.25421773
MZIP 2-3	.77338603	.3285698
MZIP 3-4	.41633729	.3956627
MZIP 4-5	.19367589	.302614
MZIP 5-6	.10671937	.165949
MZIP 6-7	.60869565	.333609

Die Deskription der Ergebnisse der *SMD-Technologie* für die zweite experimentelle Studie sind getrennt nach dem Surface-, Matching- und Deep-Structure Index und den einzelnen Messzeitpunkten Tabelle 27 zu entnehmen.

Tabelle 27: Ergebnisse der SMD-Technologie für die Messzeitpunkte der zweiten experimentellen Studie (N = 33)

Messzeitpunkt	Surface-Structure		Matching-Structure		Deep-Structure Erklärungsmodell		Deep-Structure Expertenmodell	
	M	SD	M	SD	M	SD	M	SD
MZP 1	5,8182	2,8334	2,6667	0,8165	0,0367	0,0621	0,0266	0,0477
MZP 2	3,9394	4,6767	1,6061	1,5600	0,0662	0,1052	0,0431	0,0701
MZP 3	8,3333	5,0847	3,0303	1,5509	0,1188	0,1170	0,0826	0,0852
MZP 4	10,8788	5,8671	3,8182	1,7402	0,1167	0,1147	0,0938	0,0904
MZP 5	11,0000	6,2099	3,9394	1,8361	0,1316	0,1232	0,0941	0,0923
MZP 6	11,3636	6,2190	3,9394	1,8017	0,1304	0,1215	0,0935	0,0936
MZP 7	8,4848	4,0009	3,5758	1,5417	0,1385	0,1371	0,0973	0,0875

Der Surface-Structure Index (Summe der durchschnittlich verwendeten Propositionen) nimmt zwischen dem ersten (MZP 1; $\theta = 5.81$, $SD = 2.83$) und zweiten Messzeitpunkt (MZP 2; $\theta = 3.93$, $SD = 4.67$) ab. Zum dritten Messzeitpunkt (MZP 3; $\theta = 8.33$, $SD = 5.08$) ist ein starker Anstieg der verwendeten Propositionen festzustellen. Der höchste durchschnittliche Wert des Surface-Structure Index konnte zum sechsten Messzeitpunkt (MZP 6; $\theta = 11.36$, $SD = 6.21$) diagnostiziert werden. Der erneute Abfall der durchschnittlich verwendeten Propositionen zum siebten Messzeitpunkt (MZP 7; $\theta = 8.48$, $SD = 4.0$) kann wie in der ersten Studie auf die pädagogische Intervention (*Präsentation des Lösungsmodells*) zurückgeführt werden (Tabelle 27 und Abbildung 32).

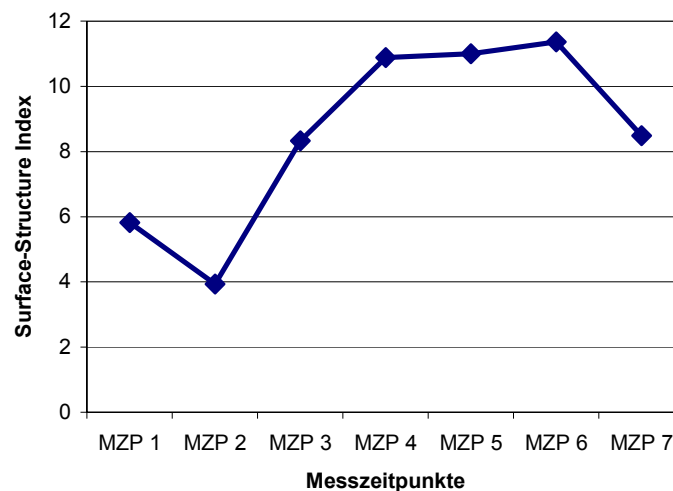


Abbildung 32: Surface-Structure Index der zweiten experimentellen Studie

Der Matching-Structure Index (durchschnittliche Komplexität der MODELLE) fällt analog zum Surface-Structure Index zwischen dem ersten (MZP 1; $\mu = 2.66$, $SD = .81$) und zweiten (MZP 2; $\mu = 1.60$, $SD = 1.56$) Messzeitpunkt ab (vgl. Tabelle 27). Zum dritten Messzeitpunkt (MZP 3; $\mu = 3.03$, $SD = 1.55$) konnte eine starke Veränderung des Matching-Structure Index gemessen werden (vgl. Abbildung 33). Die größte durchschnittliche Komplexität der Modelle konnte zum Messzeitpunkt sechs (MZP 6; $\mu = 3.93$, $SD = 1.80$) diagnostiziert werden. Zum sechsten Messzeitpunkt wurde der maximale Durchmesser der Modelle mit $Max = 8$ erreicht. Zum siebten Messzeitpunkt (MZP 7; $\mu = 3.57$, $SD = 1.54$) wird erneut ein geringerer Matching-Structure Index gemessen (vgl. Abbildung 33).

Mittels zweier Referenzmodelle (*Erklärungs-* und *Expertenmodell*) wurde die semantisch-strukturelle Ähnlichkeit (Deep-Structure) zu den einzelnen MODELLEN bestimmt. Zum ersten Messzeitpunkt weisen die MODELLE nur eine sehr geringe durchschnittliche semantisch-strukturelle Ähnlichkeit zum *Erklärungs-* ($\delta = .03$, $SD = .06$) bzw. *Expertenmodell* ($\delta = .02$, $SD = .04$) auf (vgl. Tabelle 27). Während der Arbeit mit dem multimedialen Lernprogramm ist ein Anstieg der semantisch-strukturellen Ähnlichkeit zum *Erklärungs-* und *Expertenmodell* festzustellen (vgl. Abbildung 34).

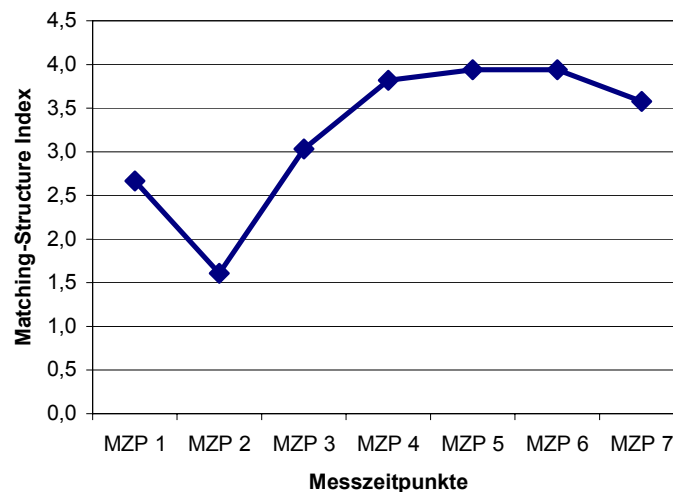


Abbildung 33: Matching-Structure Index der zweiten experimentellen Studie

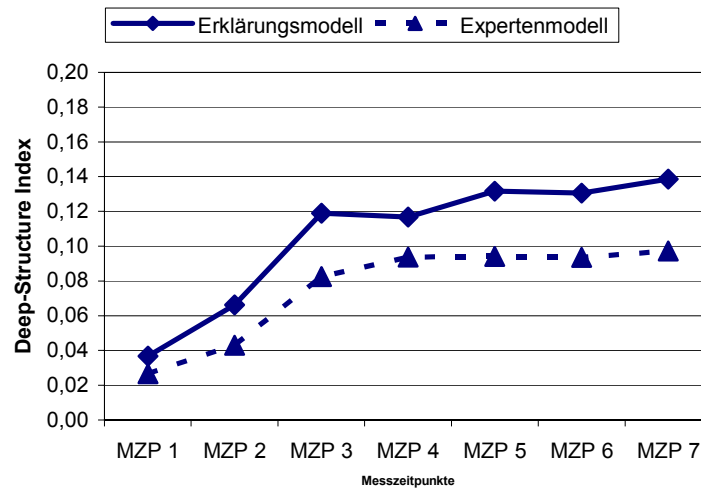


Abbildung 34: Deep-Structure Index (Erklärungs- und Expertenmodell) der zweiten experimentellen Studie

Dabei liegt der Ähnlichkeitswert des *Erklärungsmodells* während der gesamten Lernphase über dem des *Expertenmodells*. Der höchste durchschnittliche Deep-Structure Index wird am siebten Messzeitpunkt (MZIP 7; $\delta = .13$, $SD = .13$) zum *Erklärungsmodell* erreicht.

8.2.2.2 Pädagogische Interventionen

Die folgenden Darstellungen beleuchten die lernabhängige Veränderung der Modelle getrennt nach den unterschiedlichen pädagogischen Interventionen bzw. Untersuchungsgruppen *scaffolding-based learning* und *self-guided learning*.

Tabelle 28: Durchschnittliche Veränderungswahrscheinlichkeit der MODELLE getrennt nach den Untersuchungsgruppen (Oberflächen-Struktur-Vergleich; Studie 2)

Messzeitpunkt	scaffolding-based learning		self-guided learning	
	M	SD	M	SD
MZIP 1-2	.94629156	.14786557	.78804348	.31747071
MZIP 2-3	.84143223	.25856828	.70108696	.38493201
MZIP 3-4	.55242967	.34295664	.27173913	.40646942
MZIP 4-5	.18414322	.27946809	.20380435	.33442245
MZIP 5-6	.10741688	.11097357	.10597826	.21357151
MZIP 6-7	.55754476	.29816519	.66304348	.36943732

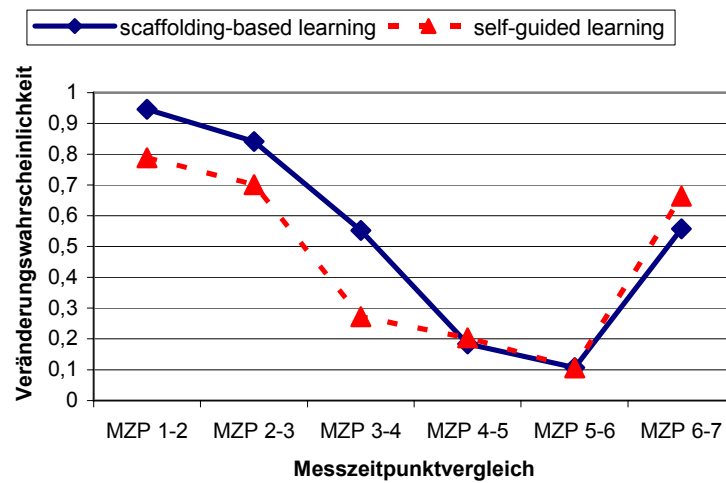


Abbildung 35: Oberflächen-Struktur-Vergleich der zweiten experimentellen Studie getrennt nach den Untersuchungsgruppen

Analog zur ersten experimentellen Studie wird in der zweiten experimentellen Studie mittels *Oberflächen-Struktur-Vergleich* eine hohe durchschnittliche Veränderungswahrscheinlichkeit zwischen dem ersten und zweiten Messzeitpunkt (MZP 1-2) diagnostiziert (vgl. Abbildung 35). Zwischen dem dritten und vierten Messzeitpunkt (MZP 3-4) sah die experimentelle Variation ein *detailliertes Feedback* (prozessorientierte Lernhilfe; vgl. Kapitel 7.1.2.2) für die Untersuchungsgruppe *scaffolding-based learning* vor. Wie den Ergebnissen in Tabelle 28 zu entnehmen ist, weisen die Probanden der Untersuchungsgruppe *scaffolding-based learning* (55 %, $SD = .34$) zwischen dem dritten und vierten Messzeitpunkt (MZP 3-4) eine höhere Veränderungswahrscheinlichkeit der MODELLE als die Untersuchungsgruppe *self-guided learning* (27 %, $SD = .40$) auf. Zwischen dem fünften und sechsten Messzeitpunkt (MZP 5-6) ist die niedrigste durchschnittliche Veränderungswahrscheinlichkeit der MODELLE bei der Untersuchungsgruppe *self-guided learning* (10 %, $SD = .21$) festzustellen. Erneut konnte zwischen dem sechsten und siebten Messzeitpunkt (MZP 6-7; pädagogische Intervention: *Präsentation des Lösungsmodells*) ein Anstieg der Veränderungswahrscheinlichkeit für beide Untersuchungsgruppen diagnostiziert werden.

Als nächster Schritt der Datenanalyse wurden die MODELLE mittels der *SMD-Technologie* getrennt nach den Untersuchungsgruppen analysiert. Die ausführlichen Ergebnistabellen für die Surface-, Matching- und Deep-Structure sind im Anhang G einzusehen.

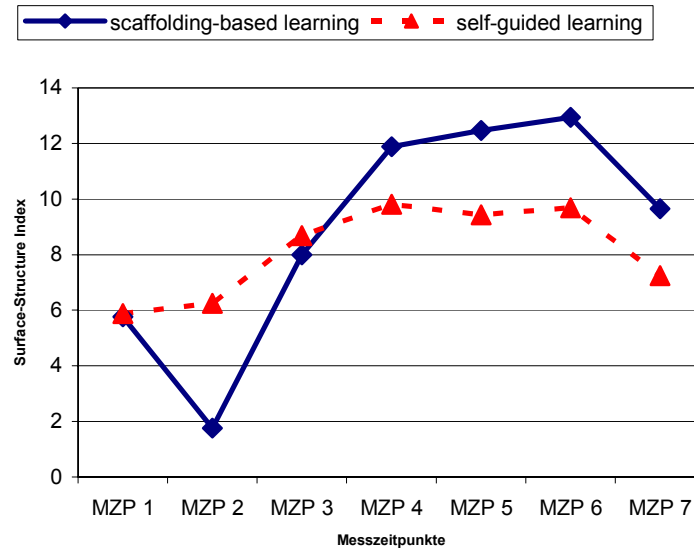


Abbildung 36: Surface-Structure Index der zweiten experimentellen Studie getrennt nach den Untersuchungsgruppen

Der Surface-Structure Index (Summe der verwendeten Propositionen) fällt bei der Untersuchungsgruppe *scaffolding-based learning* zum zweiten Messzeitpunkt stark ab (MZIP 2; $\theta_{\text{scaffolding-based learning}} = 1.76$, $SD = 2.63$). Der Unterschied am zweiten Messzeitpunkt zwischen den Untersuchungsgruppen ist signifikant ($t_{(31)} = 3.102$, $p = .004$, $\eta^2 = .23$). Die Unterschiede zwischen den Untersuchungsgruppen sind zu den anderen Messzeitpunkten nicht signifikant. Wie Abbildung 36 veranschaulicht, steigt die Summe der verwendeten Propositionen bei der Untersuchungsgruppe *scaffolding-based learning* ab dem dritten Messzeitpunkt an, bis schließlich am sechsten Messzeitpunkt (MZIP 6; $\theta_{\text{scaffolding-based learning}} = 12.94$, $SD = 6.99$) der höchste Wert diagnostiziert wurde. Auch die Probanden der Bedingung *self-guided learning* erreichen am sechsten Messzeitpunkt (MZIP 6; $\theta_{\text{self-guided learning}} = 9.69$, $SD = 4.95$) den höchsten durchschnittlichen Surface-Structure Index, dieser bleibt jedoch unter dem Wert der Untersuchungsgruppe *scaffolding-based learning*.

Die Ergebnisse der Matching-Structure Analyse (vgl. Abbildung 37) zeigen ein ähnliches Bild auf. Für die Untersuchungsgruppe *scaffolding-based learning* gilt es einen Einbruch bei der Komplexität der MODELLE zum zweiten Messzeitpunkt (MZP 2; $\mu_{\text{scaffolding-based learning}} = 1.0$, $SD = 1.5$) zu berichten. Der Unterschied zwischen den Untersuchungsgruppen ist am zweiten Messzeitpunkt signifikant ($t_{(31)} = 2.478$, $p = .019$, $\eta^2 = .16$). Ab dem dritten Messzeitpunkt weist die Untersuchungsgruppe *scaffolding-based learning* einen durchschnittlich größeren Durchmesser als die Untersuchungsgruppe *self-guided learning* auf. Die Untersuchungsgruppen unterscheiden sich jedoch im weiteren Verlauf nicht signifikant.

Die Ergebnisse der Deep-Structure Analyse werden in Abbildung 38 getrennt nach den Untersuchungsgruppen veranschaulicht. Für die Untersuchungsgruppe *scaffolding-based learning* gilt es ab dem zweiten Messzeitpunkt einen Anstieg der semantisch-strukturellen Ähnlichkeit zu den Referenzmodellen zu berichten, wobei eine größere Ähnlichkeit zum *Erklärungs-* als zum *Expertenmodell* diagnostiziert wurde. Die MODELLE der Untersuchungsgruppe *self-guided learning* sind zum siebten Messzeitpunkt ($\delta_{\text{self-guided learning}} = .16$, $SD = .14$)

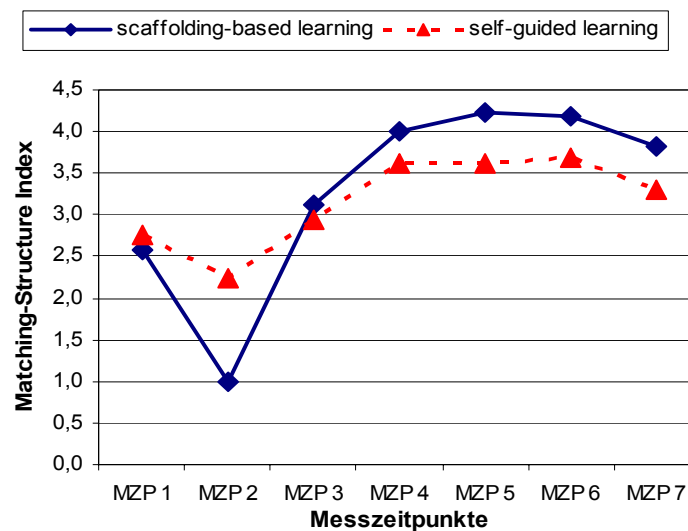


Abbildung 37: Matching-Structure Index der zweiten experimentellen Studie getrennt nach den Untersuchungsgruppen

dem *Erklärungsmodell* am ähnlichsten. Die größte Ähnlichkeit zum *Expertenmodell* erreicht die Untersuchungsgruppe *scaffolding-based learning* zum sechsten

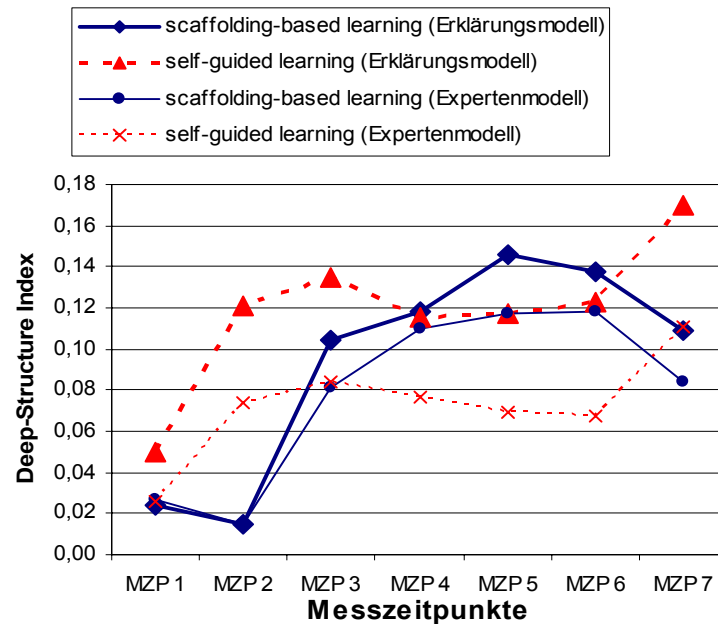


Abbildung 38: Deep-Structure Index der zweiten experimentellen Studie getrennt nach den Untersuchungsgruppen und Ähnlichkeit zum Erklärungs- bzw. Expertenmodell

Messzeitpunkt ($\delta_{\text{scaffolding-based learning}} = .11$, $SD = .11$). Die Ergebnisse der einzelnen Messzeitpunkte werden getrennt nach den Untersuchungsgruppen in Anhang G tabellarisch dargestellt. Am zweiten Messzeitpunkt (MZIP 2) unterscheiden sich die Untersuchungsgruppen bezüglich der Ähnlichkeit zum *Erklärungsmodell* ($t_{(31)} = 3.340$, $p = .002$, $\eta^2 = .26$) und zum *Expertenmodell* ($t_{(31)} = 2.642$, $p = .013$, $\eta^2 = .18$) signifikant. Zu den restlichen Messzeitpunkten gilt es keine signifikanten Gruppenunterschiede zu berichten.

8.2.2.3 Effekte von Lernervoraussetzungen

Auch für die zweite experimentelle Studie wurden die Lernervoraussetzungen *deklaratives Wissen*, *Lernstrategien* und *allgemeine kognitive Fähigkeiten* erhoben. Die Ergebnisse werden in Tabelle 29 getrennt nach den Untersuchungsgruppen dargestellt.

Tabelle 29: Skalenwerte für deklaratives Wissen, Lernstrategien und allgemeine kognitive Fähigkeiten getrennt nach den Untersuchungsgruppen (Studie 2)

	self-guided learning		scaffolding-based learning	
	M	SD	M	SD
deklaratives Wissen	8.81	3.936	7.41	3.242
<i>Lernstrategien</i>				
Organisieren (LIST)	2.71	.910	2.87	.653
Elaborieren (LIST)	3.20	.631	3.45	.482
Kritisches Prüfen (LIST)	2.91	.666	2.92	.509
Wiederholen (LIST)	3.15	.762	3.36	.681
Metakognitive Strategien (LIST)	3.46	.686	3.58	.418
<i>allgemeine kognitive Fähigkeiten</i>				
Einfallsreichtum (BIS)	87.63	11.230	91.71	7.174
sprachgebundenes Denken (BIS)	98.25	8.054	95.59	12.644
figural-bildhaftes Denken (BIS)	90.63	11.153	91.06	7.250

Im deklarativen Wissenstest *Geophysik* löste die Untersuchungsgruppe *scaffolding-based learning* durchschnittlich $M = 7.41$ ($SD = 3.242$) der insgesamt 19 Multiple-Choice-Fragen richtig. Die Untersuchungsgruppe *self-guided learning* löste $M = 8.81$ ($SD = 3.936$) Fragen richtig.

Hinsichtlich der *Lernstrategien* gaben die Probanden der Untersuchungsgruppe *scaffolding-based learning* ($M = 3.58$, $SD = .418$) für die Skala *Metakognitive Strategien* den höchsten durchschnittlichen Wert auf der 5-stufigen Ratingskala (1 = sehr selten, 2 = selten, 3 = manchmal, 4 = oft, 5 = sehr oft) des LIST-Tests an. Mit dem niedrigsten Wert wurde die Skala *Organisieren* von der Untersuchungsgruppe *scaffolding-based learning* ($M = 2.71$, $SD = .910$) eingeschätzt.

Aussagen über die *allgemeinen kognitiven Fähigkeiten* der Probanden werden durch relevante Skalen des BIS-Tests (vgl. Kapitel 7.3) gemessen. Höchster durchschnittlicher Wert erreicht die Untersuchungsgruppe *self-guided learning* für die Skala *sprachgebundenes Denken* ($M = 98.25$, $SD = 8.054$). Tiefster durchschnittlicher Wert wird ebenfalls durch die Untersuchungsgruppe *self-guided learning* in der Skala *Einfallsreichtum* ($M = 87.63$, $SD = 11.230$) erreicht.

8.2.3 Hypothesenprüfende Darstellung

Die im Hinblick auf die lernabhängige Veränderung mentaler Modelle formulierten Hypothesen werden mittels HLM-Analysen geprüft. Hierzu werden zwei Auswertungsschritte benötigt. Zunächst wird in einem ersten Wachstumsmodell (*Level-1* Modell; Gleichung 8.2.3.1) die intraindividuelle Veränderung bezüglich der abhängigen Variablen (Surface-, Matching- und Deep-Structure Indizes der *SMD-Technologie*) modelliert:

$$Y_{ti} = \pi_{0i} + \pi_{1i}(TIME) + e_{ti} \quad [8.2.3.1]$$

Dabei ist π_{1i} der Wachstumsparameter der Person i , dargestellt als linearer Zeiteffekt ($TIME$).

Die durchschnittliche Merkmalsausprägung des Surface-Structure Index liegt zum ersten Messzeitpunkt (MZP 1) mit 5.81 signifikant über 0 (vgl. Tabelle 30). Die lernabhängige Veränderung von 0.911 pro Messzeitpunkt ist ebenfalls signifikant. Das *Level-1* Modell klärt 64,9 % der Varianz auf. Somit kann $H_{1[1.1]}$ angenommen werden, welche besagt, dass sich die Anzahl der Propositionen (Surface-Structure) der MODELLE während des Lernprozesses signifikant verändern.

Tabelle 30: Lineares Wachstumsmodell der Surface-Structure (Studie 2)

Fixe Effekte	Coefficient	SE	t-Wert	df	p
Mean Initial Status π_{0i}	5.811	0.581	10.001	32	< .001
Mean Growth Rate π_{1i}	0.911	0.178	5.103	32	< .001
Zufällige Effekte	Variance Component	SD	df	χ^2	p
Initial Status r_{0i}	3.871	1.967	32	49.037	.027
Growth Rate r_{1i}	0.492	0.702	32	60.190	.002
Level-1 Error e_{ti}	3.957	15.663			

Hinsichtlich der interindividuellen Unterschiede ergeben sich sowohl im Ausgangsniveau (MZP 1; $\chi^2 = 49.037$; $p = .027$) als auch bezüglich des Wachstumsparameters ($\chi^2 = 60.190$; $p = .002$) signifikante Effekte auf der Surface-Structure Ebene.

Auch der Matching-Structure Index liegt zum ersten Messzeitpunkt (MZP 1) mit 2.33 signifikant über 0. Pro Messzeitpunkt wurde eine signifikante Ver-

änderung von 0.29 geschätzt (vgl. Tabelle 31). Das *Level-1* Modell klärt 77,5 % der Varianz auf. Folglich kann $H_{1[1.2]}$ angenommen werden; der Durchmesser (Matching-Structure) der Modelle verändert sich während des Lernprozesses. Auf interindividueller Ebene kann ein signifikanter Unterschied bezüglich der Parameter der Wachstumskurve nachgewiesen werden ($\chi^2 = 58.706$; $p = .003$).

Tabelle 31: Lineares Wachstumsmodell der Matching-Structure (Studie 2)

Fixe Effekte	Coefficient	SE	t-Wert	df	p
Mean Initial Status π_{0i}	2.335	0.165	14.154	32	< .001
Mean Growth Rate π_{1i}	0.296	0.052	5.657	32	< .001
Zufällige Effekte	Variance Component	SD	df	χ^2	p
Initial Status r_{0i}	0.255	0.505	32	44.746	.066
Growth Rate r_{1i}	0.041	0.203	32	58.706	.003
Level-1 Error e_{ij}	1.383	1.176			

Die Ähnlichkeit zum Erklärungsmodell (Deep-Structure, Erklärungsmodell) unterscheidet sich nach den Ergebnissen in Tabelle 332 zum ersten Messzeitpunkt (MZP 1) signifikant von 0. Die lernabhängige Veränderung der Deep-Structure (Erklärungsmodell) ist mit 0.015 pro Messzeitpunkt signifikant. Das *Level-1* Modell klärt 73,6 % auf. Die Hypothese $H_{1[1.3]}$ wird angenommen; die Ähnlichkeit der MODELLE zum Erklärungsmodell (Deep-Structure) verändert sich während des Lernprozesses. In Bezug auf die interindividuellen Unterschiede im Ausgangsniveau konnten keine signifikante Effekte festgestellt werden ($\chi^2 = 48.470$; $p = .031$). Für die Parameter der Wachstumskurve ergeben sich signifikante interindividuelle Unterschiede ($\chi^2 = 99.693$; $p < .001$). Die Probanden unterscheiden sich demnach nicht signifikant hinsichtlich ihres Ausgangsniveaus, jedoch hinsichtlich ihrer lernabhängigen Veränderung des Deep-Structure Index (Ähnlichkeit zum Erklärungsmodell).

Tabelle 332: Lineares Wachstumsmodell der Deep-Structure, Erklärungsmodell (Studie 2)

Fixe Effekte	<i>Coefficient</i>	<i>SE</i>	<i>t-Wert</i>	<i>df</i>	<i>p</i>
Mean Initial Status π_{0i}	0.057	0.011	5.230	32	< .001
Mean Growth Rate π_{1i}	0.015	0.004	3.635	32	.001
Zufällige Effekte	<i>Variance Component</i>	<i>SD</i>	<i>df</i>	χ^2	<i>p</i>
Initial Status r_{0i}	0.00137	0.03695	32	48.470	.031
Growth Rate r_{1i}	0.00043	0.02078	32	99.693	< .001
Level-1 Error e_{ti}	0.07559	0.00571			

Wie bereits für das Erklärungsmodell berichtet, liegt auch der Deep-Structure Index (Expertenmodell) zum ersten Messzeitpunkt (MZIP 1) mit 0.04 signifikant über 0. Außerdem wurde pro Messzeitpunkt eine signifikant positive Veränderung des Deep-Structure Index (Expertenmodell) von 0.001 festgestellt (vgl. Tabelle 323). Das *Level-1* Modell klärt 80,7 % der Varianz auf. Folglich wird Hypothese $H_{1[1.3]}$ angenommen; die Ähnlichkeit der MODELLE mit dem Expertenmodell (Deep-Structure) verändert sich während des Lernprozesses. Auf interindividueller Ebene konnte ein signifikanter Unterschied zum ersten Messzeitpunkt (MZIP 1) festgestellt werden ($\chi^2 = 49.452$; $p = .025$). Die Parameter der Wachstumskurve weisen auch signifikante interindividuelle Unterschiede auf ($\chi^2 = 83.023$; $p < .001$). Die Probanden unterscheiden sich somit auch signifikant hinsichtlich ihres Ausgangsniveaus und ihrer positiven Veränderung des Ähnlichkeitsindex zum Expertenmodell (Deep-Structure, Exper-

Tabelle 323: Lineares Wachstumsmodell der Deep-Structure, Expertenmodell (Studie 2)

Fixe Effekte	<i>Coefficient</i>	<i>SE</i>	<i>t-Wert</i>	<i>df</i>	<i>p</i>
Mean Initial Status π_{0i}	0.041	0.007	5.410	32	< .001
Mean Growth Rate π_{1i}	0.011	0.002	4.242	32	< .001
Zufällige Effekte	<i>Variance Component</i>	<i>SD</i>	<i>df</i>	χ^2	<i>p</i>
Initial Status r_{0i}	0.00067	0.02593	32	49.452	.025
Growth Rate r_{1i}	0.00015	0.01230	32	83.023	< .001
Level-1 Error e_{ti}	0.00266	0.05153			

tenmodell).

Das zweite Wachstumsmodell (*Level-2 Modell*) klärt mögliche Einflüsse von zeitstabilen Drittvariablen auf (vgl. Raudenbush & Bryk, 2002). Hierfür wird das *Level-2* Submodell nach der allgemeinen Gleichung 8.1.3.2 konstruiert. Für jede der Ebenen der *SMD-Technologie* (Surface-, Matching- und Deep-Structure) wurde eine HLM-Analyse nach Gleichung 8.1.3.2 mit den Prädiktoren *GRUPPE* (0 = *self-guided learning*; 1 = *scaffolding-based learning*) und *VORWISSEN* (0 = *niedrigeres Vorwissen*; 1 = *höheres Vorwissen*) durchgeführt. Die Ergebnisse werden in Tabellenform in Anhang G dargestellt. Hinsichtlich der Surface-Structure und der experimentellen Variation *GRUPPE* ergibt sich mittels Substitution der Ergebnisse folgendes *Level-2* Submodell (Gleichung 8.2.3.2):

$$\begin{aligned}\hat{\pi}_{0i} &= 6.883 - 2.081GRUPPE_i \\ \hat{\pi}_{1i} &= 0.419 + 0.954GRUPPE_i\end{aligned}\quad [8.2.3.2]$$

Die Probanden der Bedingung *self-guided learning* erreichen zum ersten Messzeitpunkt (MZIP 1) einen Surface-Structure Index von 6.883, welcher sich signifikant von 0 unterscheidet. Der Surface-Structure Index der Probanden der Bedingung *scaffolding-based learning* ist zum ersten Messzeitpunkt im Durchschnitt um -2.081 (4.802) niedriger. Die Untersuchungsgruppen unterscheiden sich zum ersten Messzeitpunkt (MZIP 1) jedoch nicht signifikant (vgl. Anhang G; 68,6 % aufgeklärte Varianz). Die lernabhängige Veränderung des Surface-Structure Index wird im zweiten Teil der Gleichung 8.2.3.2 angegeben. Demnach verändert sich der Surface-Structure Index der Probanden unter der Bedingung *self-guided learning* um 0.419; für Probanden der Bedingung *scaffolding-based learning* wurde die Veränderung des Surface-Structure Index pro Messzeitpunkt signifikant höher geschätzt (0.419+0.954=1.373). Somit wird die Hypothese $H_{1[2,1]}$ angenommen; die lernabhängige Veränderung der Anzahl der Propositionen (Surface-Structure) der MODELLE unterscheidet sich zwischen den Untersuchungsgruppen *self-guided learning* und *scaffolding-based learning*.

Die Ergebnisse bezüglich des Prädiktors *VORWISSEN* zeigen, dass sich die Probanden mit *niedrigerem Vorwissen* zum ersten Messzeitpunkt (MZIP 1) signifikant von 0 unterscheiden. Der Unterschied zu Probanden mit *höherem Vorwissen* ist nicht signifikant. Die Hypothese $H_{0[3,1]}$ wird beibehalten, welche besagt, dass das bereichsspezifische Vorwissen keinen Einfluss auf die lernab-

hängige Veränderung der Anzahl der Propositionen (Surface-Structure) hat (vgl. Anhang G; 61,7 % aufgeklärte Varianz).

Für die Ebene der Matching-Structure wurde eine weitere HLM-Analyse durchgeführt. Aus der allgemeinen Formel des *Level-2* Submodells ergibt sich durch Substitution folgende Gleichung (8.2.3.3) für den Prädiktor *GRUPPE*:

$$\begin{aligned}\hat{\pi}_{0i} &= 2.607 - 0.527GRUPPE_i \\ \hat{\pi}_{1i} &= 0.294 + 0.038GRUPPE_i\end{aligned}\quad [8.2.3.3]$$

Sowohl die Probanden der Bedingung *self-guided learning* als auch die Probanden der Bedingung *scaffolding-based learning* verändern den Durchmesser der MODELLE (Matching-Structure) signifikant während der Lernphase. Den Ergebnissen in Anhang G ist weiter zu entnehmen, dass die Probanden der Bedingung *scaffolding-based learning* eine signifikant größere lernabhängige Veränderung aufweisen (pro MZP 0.332; 74,1 % aufgeklärte Varianz). Folglich wird die Hypothese $H_{1[2.2]}$ angenommen; die lernabhängige Veränderung des Durchmessers (Matching-Structure) der MODELLE unterscheidet sich zwischen den Untersuchungsgruppen.

Für den Prädiktor *VORWISSEN* konnte ein signifikanter Effekt für die Untersuchungsgruppe mit *niedrigerem Vorwissen* zum ersten Messzeitpunkt (MZP 1) und hinsichtlich deren lernabhängigen Veränderung festgestellt werden (vgl. Anhang G; 76,7 % aufgeklärte Varianz). Die Probanden mit *höherem Vorwissen* unterscheiden sich von den Probanden mit *niedrigerem Vorwissen* jedoch nicht bezüglich dem Wert zum ersten Messzeitpunkt und deren lernabhängigen Veränderung der Matching-Structure. Somit wird die Hypothese $H_{0[3.2]}$ beibehalten. Das bereichsspezifische Vorwissen hat keinen Einfluss auf die lernabhängige Veränderung des Durchmesser (Matching-Structure).

In einer weiteren HLM-Analyse wurde im Hinblick auf die Ebene der Deep-Structure (Erklärungsmodell) folgendes *Level-2* Submodell ermittelt (Gleichung 8.2.3.4):

$$\begin{aligned}\hat{\pi}_{0i} &= 0.081 - 0.046GRUPPE_i \\ \hat{\pi}_{1i} &= 0.012 + 0.007GRUPPE_i\end{aligned}\quad [8.2.3.4]$$

Die Ergebnisse in Anhang G zeigen keine signifikanten Unterschiede für die Untersuchungsgruppen (*self-guided learning* und *scaffolding-based learning*) bezüglich deren lernabhängigen Veränderung (82,3 % aufgeklärte Varianz). Die Hypothese $H_{0[2,3]}$ wird beibehalten; die lernabhängige Veränderung der Ähnlichkeit der MODELLE mit dem Erklärungsmodell (Deep-Structure) unterscheidet sich nicht zwischen den Untersuchungsgruppen *self-guided learning* und *scaffolding-based learning*.

Die Ergebnisse für den Prädiktor *VORWISSEN* zeigen, dass sich die Probanden mit *niedrigerem* und *höherem Vorwissen* bezüglich deren lernabhängigen Veränderung nicht signifikant unterscheiden (vgl. Anhang G; 72,8 % aufgeklärte Varianz). Die Hypothese $H_{0[3,3]}$ wird beibehalten, welche besagt, dass das bereichsspezifische Vorwissen in keinem Zusammenhang mit der lernabhängigen Veränderung der Ähnlichkeit eines konstruierten MODELLS zum Erklärungsmodell (Deep-Structure) steht.

Die folgende HLM-Analyse (Prädiktor *GRUPPE*) ergab für die Deep-Structure (Expertenmodell) in Gleichung 8.2.3.5 dargestelltes *Level-2* Submodell:

$$\begin{aligned}\hat{\pi}_{0i} &= 0.048 - 0.013GRUPPE_i \\ \hat{\pi}_{1i} &= 0.008 + 0.006GRUPPE_i\end{aligned}\quad [8.2.3.5]$$

Die Ergebnisdarstellung in Anhang G zeigt keine signifikanten Unterschiede für die Untersuchungsgruppen (*self-guided learning* und *scaffolding-based learning*) bezüglich der lernabhängigen Veränderung auf (80,1 % aufgeklärte Varianz). Die Hypothese $H_{0[2,3]}$ wird beibehalten; die lernabhängige Veränderung der Ähnlichkeit der MODELLE mit dem Expertenmodell (Deep-Structure) unterscheidet sich nicht zwischen den Untersuchungsgruppen.

Für den Prädiktor *VORWISSEN* kann ein signifikanter Unterschied bezüglich der lernabhängigen Veränderung (Deep-Structure, Expertenmodell) zwischen Probanden mit *niedrigerem Vorwissen* (0.003) und Probanden mit *höherem Vorwissen* (0.011) berichtet werden (vgl. Anhang G; 86,7 % aufgeklärte Varianz). Somit wird die Hypothese $H_{1[3,3]}$ angenommen: Das bereichsspezifische Vorwissen steht in Zusammenhang mit der lernabhängigen Veränderung der Ähnlichkeit eines konstruierten MODELLS zum Expertenmodell (Deep-Structure).

8.3 ERGEBNISSE DER DRITTEN STUDIE

An der dritten experimentellen Studie nahmen 47 Probanden teil (vgl. Kapitel 7.2.3). Die Inhaltsdomäne und der Themenbereich bzw. die Problemstellung (Geophysik - Plattentektonik) wie auch die acht Messzeitpunkte (vgl. Kapitel 7.1.2.3) wurden analog zur zweiten Studie beibehalten. Daran anschließend konnte das *Erklärungs-* bzw. *Expertenmodell* für die dritte Studie ohne Veränderungen übernommen werden (vgl. Anhang D). Die experimentelle Variation der Untersuchungsgruppen sah im Gegensatz zur ersten und zweiten Studie eine zusätzliche *Kontrollgruppe* vor.

8.3.1 Vorbetrachtung der Untersuchungsstichprobe

Die Stichprobe der dritten Studie wurde bereits in Kapitel 7.2.3 dargestellt. Die Ergebnisse des Chi-Quadrat-Tests zeigen, dass sich die Untersuchungsstichprobe innerhalb der experimentellen Variation hinsichtlich der geschlechtsspezifischen Verteilung nicht unterscheiden (vgl. Tabelle 34).

Tabelle 34: Absolute Häufigkeiten und Prozentangaben der Untersuchungsstichprobe für die experimentellen Bedingungen *self-guided learning* (N = 18), *scaffolding-based learning* (N = 19) und *Kontrollgruppe* (N = 10), Ergebnisse des Chi-Quadrat-Tests (Studie 3)

	self-guided learning		scaffolding-based learning		Kontrollgruppe		χ^2	df	p
	N	%	N	%	N	%			
Geschlecht									
weiblich	8	61.5	9	69.2	9	69.2			
männlich	5	38.5	4	30.8	4	30.8	1.915	2	.384

Auf Grund der explorativen Datenanalyse mittels *Shapiro-Wilk-Test* wird die Annahme der Normalverteilung für den Faktor *Alter* ($W = .877$, $df = 47$, $p = .001$) zurückgewiesen. Für den Faktor *deklarativer Wissenstest* kann von einer normalverteilten Datengrundlage ausgegangen werden.

Tabelle 35: Mittelwerte und Standardabweichungen der Untersuchungsstichprobe für die experimentellen Bedingungen *self-guided learning* (N = 18), *scaffolding-based learning* (N = 19) und Kontrollgruppe (N = 10), Ergebnisse des Kruskal-Wallis-H-Tests (Studie 3)

	self-guided learning		scaffolding-based learning		Kontrollgruppe				
	N	Mittlerer Rang	N	Mittlerer Rang	N	Mittlerer Rang	H	df	p
Alter	18	24.25	19	21.00	10	29.25	2.456	2	.293

Die Ergebnisse des *Kruskal-Wallis-H-Tests* zeigen, dass sich die Untersuchungsgruppen hinsichtlich des Faktors *Alter* nicht signifikant unterscheiden (vgl. Tabelle 35). Hinsichtlich des Faktors *deklaratives Wissen* (Vorwissen im Bereich Geographie / Geophysik) zeigt die Analyse mittels *ANOVA* ebenfalls keine signifikanten Gruppenunterschiede (vgl. Tabelle 36). Weitere Datenanalysen und Interpretation hinsichtlich der Untersuchungsgruppen sind somit unbedenklich.

Tabelle 36: Mittelwerte und Standardabweichungen der Untersuchungsstichprobe für die experimentellen Bedingungen *self-guided learning* (N = 18), *scaffolding-based learning* (N = 19) und Kontrollgruppe (N = 10), Ergebnisse der ANOVA (Studie 3)

		Quadrat-summe	df	Mittel der Quadrate	F	p
deklaratives Wissen	Zwischen den Gruppen	14.048	2	7.024	.837	.440
	Innerhalb der Gruppen	369.058	44	8.388		
	Gesamt	383.106	46			

8.3.2 Deskription der Ergebnisse

Die Deskription der Ergebnisse berücksichtigt die in Kapitel 6 formulierten Fragestellungen. Zunächst werden die Ergebnisse des *Oberflächen-Struktur-Vergleiches* und anschließend die der *SMD-Technologie* dargestellt.

8.3.2.1 Diagnose lernabhängiger Veränderung

Für den *Oberflächen-Struktur-Vergleich* (vgl. Kapitel 3.2.5 und 7.3.6) konnten 17 *Model-Rater* gewonnen werden. Diese begutachteten die in chronologischer Reihenfolge projizierten MODELLE bezüglich deren Ähnlichkeiten.

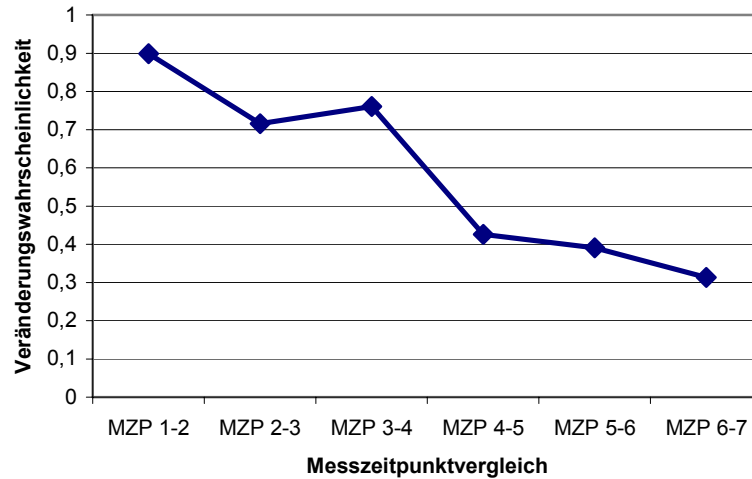


Abbildung 39: Oberflächen-Struktur-Vergleich der dritten experimentellen Studie

Zwischen dem ersten und zweiten Messzeitpunkt (MZIP 1-2; 89 %, $SD = .226$) wurde, wie in den beiden Studien zuvor, eine hohe durchschnittliche Veränderungswahrscheinlichkeit festgestellt (vgl. Abbildung 39). Während der Lernphase kann bis auf den Vergleich zwischen Messzeitpunkt drei und vier (MZIP 3-4; 76 %, $SD = .349$) eine Abnahme der Veränderungswahrscheinlichkeit berichtet werden. Die niedrigste durchschnittliche Veränderungswahrscheinlichkeit kann trotz der pädagogischen Intervention *Präsentation des Lösungsmodells* zwischen dem sechsten und siebten Messzeitpunkt (MZIP 6-7; 31 %, $SD = .348$) festgestellt werden (vgl. Tabelle 37).

Tabelle 37: Durchschnittliche Veränderungswahrscheinlichkeit der MODELLE (N = 47) nach dem Oberflächen-Struktur-Vergleich (Studie 3)

Messzeitpunkt	Veränderungswahrscheinlichkeit	
	M	SD
MZIP 1-2	.89862328	.22639904
MZIP 2-3	.71589487	.37631853
MZIP 3-4	.76095119	.34917088
MZIP 4-5	.42553191	.42012212
MZIP 5-6	.39048811	.38853463
MZIP 6-7	.31289111	.34802769

Die Ergebnisse der anschließenden Analyse mittels *SMD-Technologie* werden in Tabelle 38 veranschaulicht. Die umfassende Analyse zeigt für jeden Messzeitpunkt die Ausprägungen des Surface-, Matching- und Deep-Structure Index.

Tabelle 38: Ergebnisse der SMD-Technologie für die Messzeitpunkte der dritten experimentellen Studie (N = 47)

Messzeitpunkt	Surface-Structure		Matching-Structure		Deep-Structure Erklärungsmodell		Deep-Structure Expertenmodell	
	M	SD	M	SD	M	SD	M	SD
MZP 1	3,7660	2,9797	1,7660	1,2548	0,0803	0,1043	0,0700	0,0884
MZP 2	3,4468	3,8096	1,3191	1,4158	0,0810	0,1198	0,0581	0,0877
MZP 3	4,4468	4,4272	1,9787	1,7753	0,0938	0,1375	0,0590	0,0734
MZP 4	4,8085	3,9212	2,1702	1,5368	0,0882	0,1037	0,0582	0,0755
MZP 5	7,6170	5,3631	2,9574	1,6545	0,1186	0,1259	0,0859	0,0882
MZP 6	8,0000	6,0828	2,9149	1,4269	0,0980	0,0922	0,0784	0,0841
MZP 7	6,5532	4,4663	2,8936	1,5070	0,1138	0,1070	0,0679	0,0691

Die Summe der verwendeten Propositionen (Surface-Structure) zeigt nur eine geringe Veränderung vom ersten (MZP 1; $\theta = 3.77$, $SD = 2.98$) bis zum vierten (MZP 4; $\theta = 4.81$, $SD = 3.92$) Messzeitpunkt. Zum fünften Messzeitpunkt (MZP 5; $\theta = 7.62$, $SD = 5.36$) kann eine Zunahme des Surface-Structure Index festgestellt werden (vgl. Abbildung 40). Der höchste durchschnittliche Surface-Structure Index wurde zum Messzeitpunkt sechs (MZP 6; $\theta = 8.00$, $SD = 6.08$) diagnostiziert.

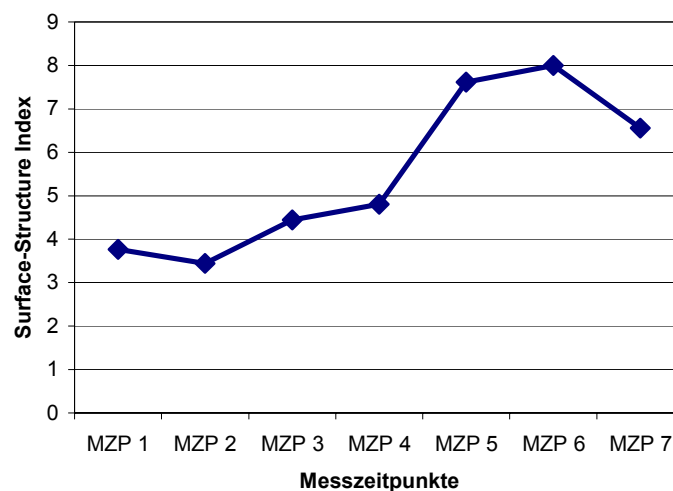


Abbildung 40: Surface-Structure Index der dritten experimentellen Studie

Abbildung 41 illustriert die lernabhängige Veränderung der MODELLE, gemessen auf der Ebene der Matching-Structure. Die niedrigste Komplexität der MODELLE wird am zweiten Messzeitpunkt (MZIP 2; $\mu = 1.32$, $SD = 1.42$) erreicht. Nach dem zweiten Messzeitpunkt steigt der Durchmesser der MODELLE an und erreicht zum fünften Messzeitpunkt (MZIP 5; $\mu = 2.96$, $SD = 1.43$) den höchsten Wert. Bis zum Ende der Lernphase am siebten Messzeitpunkt (MZIP 7; $\mu = 2.89$, $SD = 1.51$) bleibt der Matching-Structure Index nahezu konstant.

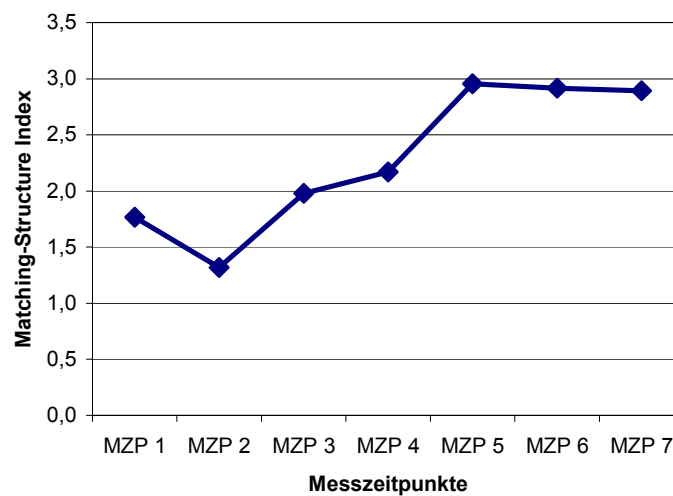


Abbildung 41: Matching-Structure Index der dritten experimentellen Studie

Die semantisch-strukturelle Ähnlichkeit der MODELLE zum *Erklärungs-* und einem *Expertenmodell* wird auf der Ebene der Deep-Structure analysiert. Wie die Ergebnisse in Tabelle 38 zeigen, weisen die MODELLE zum ersten Messzeitpunkt (MZIP 1) nur eine geringe semantisch-strukturelle Ähnlichkeit zum *Erklärungs-* ($\delta = .08$, $SD = .10$) bzw. *Expertenmodell* ($\delta = .07$, $SD = .09$) auf. Der höchste Ähnlichkeitswert wird am fünften Messzeitpunkt (MZIP 5; $\delta = .12$, $SD = .13$) zwischen den MODELLEN und dem *Erklärungsmodell* erreicht. Die lernabhängige Veränderung der MODELLE, diagnostiziert an der semantisch-strukturellen Ähnlichkeit zum *Erklärungs-* bzw. *Expertenmodell* wird in Abbildung 42 veranschaulicht.

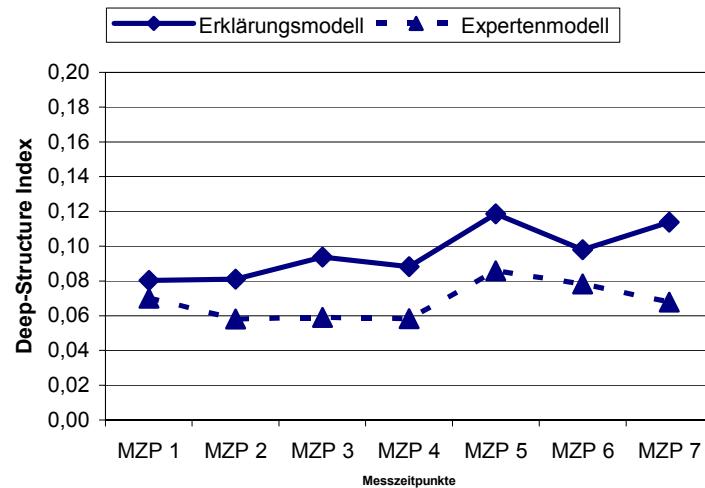


Abbildung 42: Deep-Structure Index (Erklärungs- und Expertenmodell) der dritten experimentellen Studie

8.3.2.2 Pädagogische Interventionen

Die folgenden Darstellungen beleuchten die lernabhängige Veränderung der Modelle getrennt nach den unterschiedlichen pädagogischen Interventio-

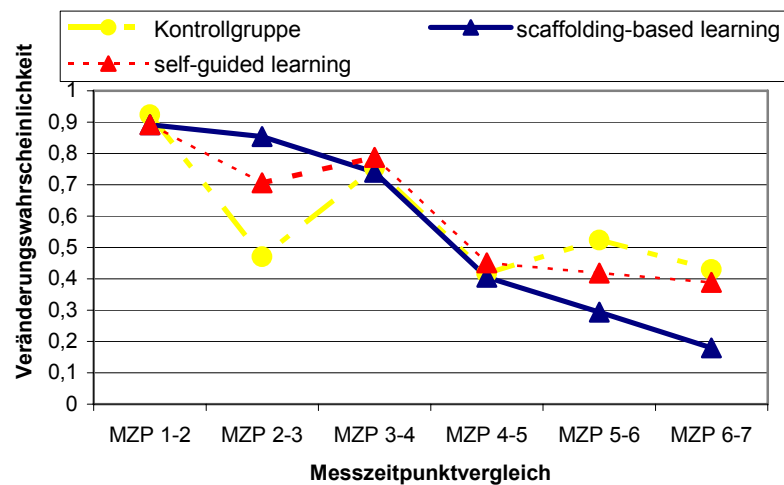


Abbildung 43: Oberflächen-Struktur-Vergleich der dritten experimentellen Studie getrennt nach den Untersuchungsgruppen

nen bzw. Untersuchungsgruppen. Im Gegensatz zur ersten und zweiten experimentellen Studie wird in der dritten Studie zwischen den Untersuchungsgruppen *scaffolding-based learning*, *self-guided learning* und *Kontrollgruppe* unterschieden.

Für die drei Untersuchungsgruppen konnte mittels *Oberflächen-Struktur-Vergleich* eine hohe durchschnittliche Veränderungswahrscheinlichkeit zwischen dem ersten und zweiten Messzeitpunkt (MZIP 1-2) diagnostiziert werden (vgl. Abbildung 43).

Tabelle 39: Durchschnittliche Veränderungswahrscheinlichkeit der MODELLE getrennt nach den Untersuchungsgruppen (Oberflächen-Struktur-Vergleich; Studie 3)

Messzeitpunkt	scaffolding-based learning		self-guided learning		Kontrollgruppe	
	M	SD	M	SD	M	SD
MZIP 1-2	.89164087	.10835913	.89215686	.10784314	.92352941	.14686261
MZIP 2-3	.85448916	.14551084	.70588235	.29411765	.47058824	.46566047
MZIP 3-4	.73993808	.26006192	.7875817	.2124183	.75294118	.40355971
MZIP 4-5	.40557276	.59442724	.45098039	.54901961	.41764706	.49406318
MZIP 5-6	.29411765	.70588235	.41830065	.58169935	.52352941	.43177343
MZIP 6-7	.17956656	.82043344	.3888889	.61111111	.42941176	.33049542

Zwischen den Messzeitpunkten vier und fünf (MZIP 4-5) sah die experimentelle Variation ein *detailliertes Feedback* (prozessorientierte Lernhilfe; vgl. Kapitel 7.1.2.2) für die Untersuchungsgruppe *scaffolding-based learning* vor. Die Untersuchungsgruppen *self-guided learning* und *Kontrollgruppe* erhielten keine zusätzliche Lernhilfe. Tabelle 39 ist zu entnehmen, dass sich die Untersuchungsgruppen zwischen dem vierten und fünften Messzeitpunkt (MZIP 4-5) nur unwesentlich unterscheiden. Die MODELLE der Probanden *scaffolding-based learning* weisen zwischen den Messzeitpunkten sechs und sieben (18 %, $SD = .82$) die niedrigste durchschnittliche Veränderungswahrscheinlichkeit auf.

In einem weiteren Schritt erfolgte die umfangreiche Analyse der MODELLE mittels der automatisierten *SMD-Technologie* getrennt nach den Untersuchungsgruppen *scaffolding-based learning*, *self-guided learning* und *Kontrollgruppe*. Eine detaillierte Ergebnisdarstellung der Surface-, Matching- und Deep-Structure erfolgt in Anhang H.

Der höchste Ausprägung des Surface-Structure Index (Summe der verwendeten Propositionen) der Untersuchungsgruppe *scaffolding-based learning* wird am fünften Messzeitpunkt (MZIP 5; $\theta_{\text{scaffolding-based learning}} = 8.33$, $SD = 4.31$) ge-

messen. Der Unterschied der Probanden in den einzelnen Versuchsbedingungen am fünften Messzeitpunkt (MZIP 5) ist signifikant ($F_{(2,44)} = 3.657, p = .034, \eta^2 = .14$). Der *a posteriori*-Test (Scheffé-Test) zeigt, dass der signifikante Effekt zwischen den Untersuchungsgruppen *Kontrollgruppe* und *scaffolding-based learning* liegt ($F_{(2,44)} = 5.331, p = .035$). Abbildung 44 veranschaulicht den höchsten Surface-Structure Index der Untersuchungsgruppe *self-guided learning* am sechsten Messzeitpunkt (MZIP 6; $\theta_{\text{self-guided learning}} = 8.78, SD = 7.88$) und der Untersuchungsgruppe *Kontrollgruppe* am siebten Messzeitpunkt (MZIP 7; $\theta_{\text{Kontrollgruppe}} = 7.40, SD = 4.32$). Die Untersuchungsgruppen unterscheiden sich am siebten Messzeitpunkt (MZIP 7) signifikant ($F_{(2,44)} = 3.686, p = .033, \eta^2 = .14$). Der *a posteriori*-Test (Scheffé-Test) zeigt, dass der signifikante Effekt der ANOVA durch die Unterschiede zwischen den Untersuchungsgruppen *Kontrollgruppe* und *scaffolding-based learning* hervorgerufen wird ($F_{(2,44)} = 4.173, p = .051$).

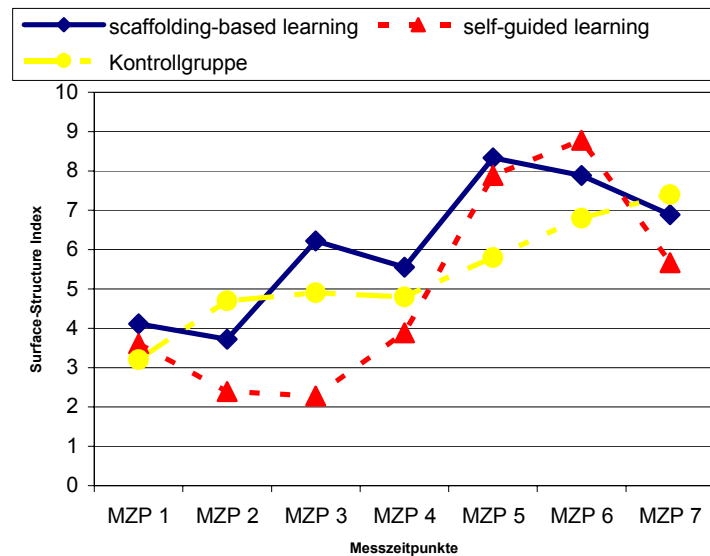


Abbildung 44: Surface-Structure Index der dritten experimentellen Studie getrennt nach den Untersuchungsgruppen

Die Komplexität der MODELLE in Bezug auf deren lernabhängige Veränderung wird in Abbildung 45 getrennt nach den Untersuchungsgruppen dargestellt. Der durchschnittlich größte Durchmesser der Modelle konnte für die Untersuchungsgruppe *Kontrollgruppe* am siebten Messzeitpunkt (MZIP 7; $\mu_{\text{Kontrollgruppe}} = 3.4, SD = 1.5$) festgestellt werden (vgl. Abbildung 45). Die Untersu-

chungsgruppen unterscheiden sich am fünften Messzeitpunkt signifikant ($F_{(2,44)} = 5.450, p = .008, \eta^2 = .19$). Der *a posteriori*-Test (Scheffé-Test) zeigt, dass der signifikante Effekt der ANOVA durch die Unterschiede zwischen den Untersuchungsgruppen *Kontrollgruppe* und *scaffolding-based learning* hervorgerufen wird ($F_{(2,44)} = 1.926, p = .009$).

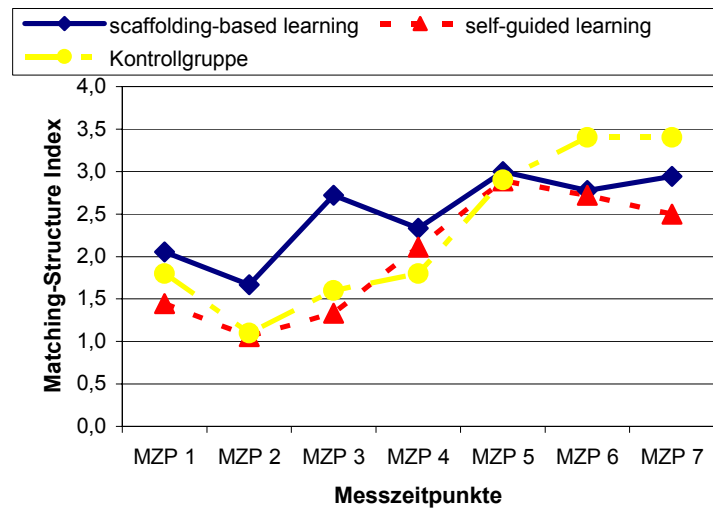


Abbildung 45: Matching-Structure Index der dritten experimentellen Studie getrennt nach den Untersuchungsgruppen

Die lernabhängige Veränderung der semantisch-strukturellen Ähnlichkeit zu den Referenzmodellen wird in Abbildung 46 veranschaulicht. Die MODELLE der Untersuchungsgruppe *scaffolding-based learning* sind am fünften Messzeitpunkt ($\delta_{\text{scaffolding-based learning}} = .17, SD = .14$) dem *Erklärungsmodell* am ähnlichsten. Die Untersuchungsgruppen unterscheiden sich am fünften Messzeitpunkt hinsichtlich dem *Erklärungsmodell* signifikant ($F_{(2,44)} = 5.034, p = .011, \eta^2 = .18$). Der *a posteriori*-Test (Scheffé-Test) zeigt, dass der signifikante Effekt der ANOVA auf die Unterschiede zwischen den Untersuchungsgruppen *Kontrollgruppe* und *self-guided learning* zurückzuführen ist ($F_{(2,44)} = 0.144, p = .011$). Auffällig bei der Untersuchungsgruppe *scaffolding-based learning* ist der starke Anstieg des Deep-Structure Index vom vierten zum fünften Messzeitpunkt und die starke Abnahme der semantisch-strukturellen Ähnlichkeit vom fünften zum sechsten Messzeitpunkt.

Die größte semantisch-strukturelle Ähnlichkeit zum *Expertenmodell* erreicht die Untersuchungsgruppe *Kontrollgruppe* am zweiten Messzeitpunkt ($\delta_{\text{Kontrollgruppe}} = .11$, $SD = .12$). Die Ergebnisse der einzelnen Messzeitpunkte werden getrennt nach den Untersuchungsgruppen in Anhang H tabellarisch dargestellt. Zwischen den Untersuchungsgruppen konnten keine signifikanten Effekte hinsichtlich der Ähnlichkeit zum *Expertenmodell* festgestellt werden.

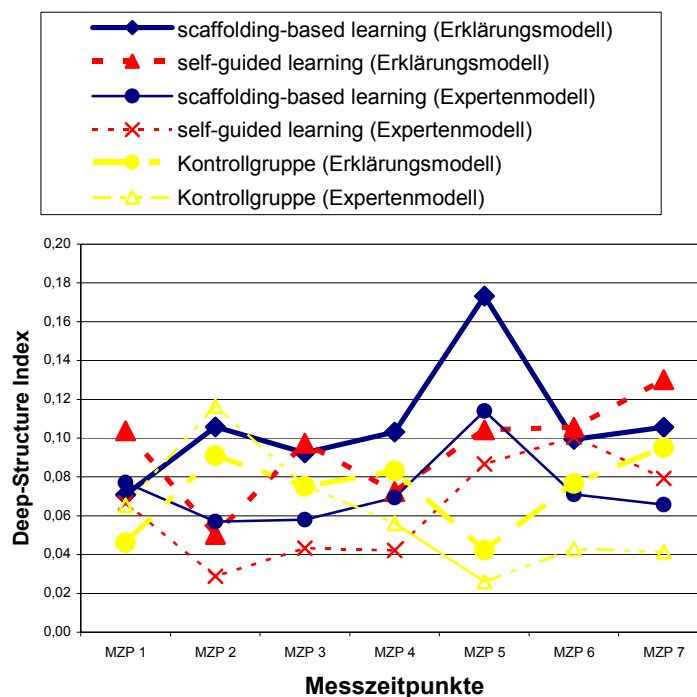


Abbildung 46: Deep-Structure Index der dritten experimentellen Studie getrennt nach den Untersuchungsgruppen und Ähnlichkeit zum Erklärungs- bzw. Expertenmodell

8.3.2.3 Effekte von Lernervoraussetzungen

Wie in der ersten und zweiten experimentellen Studie, wurden auch für die dritte Studie die Lernervoraussetzungen *deklaratives Wissen*, *Lernstrategien* und *allgemeine kognitive Fähigkeiten* erhoben. Tabelle 40 veranschaulicht die Ergebnisse getrennt nach den Untersuchungsgruppen.

Tabelle 40: Skalenwerte für deklaratives Wissen, Lernstrategien und allgemeine kognitive Fähigkeiten getrennt nach den Untersuchungsgruppen (Studie 3)

	self-guided learning		scaffolding-based learning		Kontrollgruppe	
	M	SD	M	SD	M	SD
deklaratives Wissen	10.28	2.824	10.95	2.738	9.50	3.308
<i>Lernstrategien</i>						
Organisieren (LIST)	3.46	.554	3.52	.841	3.67	.649
Elaborieren (LIST)	3.42	.464	3.81	.475	3.73	.579
Kritisches Prüfen (LIST)	3.13	.553	3.21	.591	3.24	.563
Wiederholen (LIST)	2.85	.522	2.71	.806	2.47	.756
Metakognitive Strategien (LIST)	3.36	.343	3.53	.507	3.33	.667
<i>allgemeine kognitive Fähigkeiten</i>						
Einfallsreichtum (BIS)	1179.17	86.286	1187.26	69.187	1134.50	127.096
sprachgebundenes Denken (BIS)	1639.89	103.388	1638.21	88.145	1616.50	114.094
figural-bildhaftes Denken (BIS)	1448.78	113.956	1471.05	86.933	1443.88	119.906

Die Untersuchungsgruppe *scaffolding-based learning* löste beim deklarativen Wissenstest durchschnittlich $M = 10.95$ ($SD = 2.738$) der insgesamt 19 Multiple-Choice-Fragen richtig. Die Untersuchungsgruppe *self-guided learning* löste $M = 10.28$ ($SD = 2.824$) und die Untersuchungsgruppe *Kontrollgruppe* löste $M = 9.50$ ($SD = 3.308$) Fragen richtig.

Die Probanden der Untersuchungsgruppe *Kontrollgruppe* gaben für die Lernstrategien-Skala *Elaborieren* ($M = 3.73$, $SD = .579$) den höchsten durchschnittlichen Wert auf der 5-stufigen Ratingskala (1 = sehr selten, 2 = selten, 3 = manchmal, 4 = oft, 5 = sehr oft) des LIST-Tests an. Mit dem niedrigsten Wert wurde die Skala *Wiederholen* ($M = 2.47$, $SD = .756$), ebenfalls von der Untersuchungsgruppe *Kontrollgruppe* eingeschätzt.

Aussagen über die *allgemeinen kognitiven* Fähigkeiten der Probanden werden durch relevante Skalen des BIS-Tests (vgl. Kapitel 7.3) gemessen. Bei den in Tabelle 40 dargestellten Werten handelt es sich um Rohwerte des BIS-Tests, da Standardwerte für die Altersgruppe der Stichprobe nicht vorliegen. Höchster durchschnittlicher Wert erreicht die Untersuchungsgruppe *self-guided learning* für die Skala *sprachgebundenes Denken* ($M = 1639.89$, $SD = 103.388$). Tiefster durch-

schnittlicher Wert wird von der Untersuchungsgruppe *Kontrollgruppe* in der Skala *Einfallsreichtum* ($M = 1134.50$, $SD = 127.096$) erreicht.

8.3.3 Hypothesenprüfende Darstellung

Ähnlich wie für die Auswertung der ersten und zweiten experimentellen Studie werden mittels HLM-Analyse die formulierten Hypothesen (vgl. Kapitel 6) für die dritte Studie geprüft. Das erste Wachstumsmodell (*Level-1* Modell; Gleichung 8.3.3.1) modelliert die intraindividuelle Veränderung der abhängigen Variable (Surface-, Matching- und Deep-Structure Indizes der *SMD-Technologie*):

$$Y_{ti} = \pi_{0i} + \pi_{1i}(TIME) + e_{ti} \quad [8.3.3.1]$$

Dabei ist π_{1i} der Wachstumsparameter der Person i , dargestellt als linearer Zeiteffekt ($TIME$).

Die Ergebnisse in Tabelle 41 verdeutlichen, dass die durchschnittliche Merkmalsausprägung des Surface-Structure Index zum ersten Messzeitpunkt (MZIP 1) mit 3.30 signifikant über 0 liegt. Auch die lernabhängige Veränderung des Surface-Structure Index pro Messzeitpunkt (0.73) ist in der dritten Studie signifikant. Das *Level-1* Modell klärt 53,3 % der Varianz auf. Die Hypothese $H_{1[1.1]}$ wird angenommen; die Anzahl der Propositionen (Surface-Structure) der MODELLE verändern sich während des Lernprozesses signifikant.

Hinsichtlich der interindividuellen Unterschiede ergeben sich im Ausgangsniveau keine signifikanten Effekte (MZIP 1; $\chi^2 = 58.757$; $p = .098$). Bezüglich des Wachstumsparameters ($\chi^2 = 84.152$; $p = .001$) konnten signifikante Effekte auf der Surface-Structure Ebene festgestellt werden (vgl. Tabelle 41).

Tabelle 41: Lineares Wachstumsmodell der Surface-Structure (Studie 3)

Fixe Effekte	Coefficient	SE	t-Wert	df	p
Mean Initial Status π_{0i}	3.308	0.430	7.680	46	< .001
Mean Growth Rate π_{1i}	0.737	0.142	5.155	46	< .001
Zufällige Effekte	Variance Component	SD	df	χ^2	p
Initial Status r_{0i}	1.894	1.376	46	58.757	.098
Growth Rate r_{1i}	0.435	0.435	46	84.152	.001
Level-1 Error e_{ti}	14.709	3.835			

Der Matching-Structure Index liegt in der dritten experimentellen Studie zum ersten Messzeitpunkt mit 1.47 signifikant über 0. Die signifikante Veränderung während des Lernprozesses wird auf 0.26 geschätzt (vgl. Tabelle 42; 26,2 % aufgeklärte Varianz). Folglich kann $H_{1[1,2]}$ angenommen werden; der Durchmesser (Matching-Structure) der MODELLE verändert sich während des Lernprozesses. Auf interindividueller Ebene konnte sowohl ein signifikanter Unterschied bezüglich des Ausgangsniveaus ($\chi^2 = 90.223$; $p < .001$) als auch bezüglich des Wachstumsparameters nachgewiesen werden ($\chi^2 = 78.545$; $p = .002$).

Tabelle 42: Lineares Wachstumsmodell der Matching-Structure (Studie 3)

Fixe Effekte	<i>Coefficient</i>	<i>SE</i>	<i>t-Wert</i>	<i>df</i>	<i>p</i>
Mean Initial Status π_{0i}	1.476	0.161	9.162	46	< .001
Mean Growth Rate π_{1i}	0.269	0.041	6.469	46	< .001
Zufällige Effekte	<i>Variance Component</i>	<i>SD</i>	<i>df</i>	χ^2	<i>p</i>
Initial Status r_{0i}	0.598	0.773	46	90.223	< .001
Growth Rate r_{1i}	0.033	0.184	46	78.545	.002
Level-1 Error e_{ti}	1.340	1.157			

Der Deep-Structure Index (Ähnlichkeit zum Erklärungsmodell) unterscheidet sich zum ersten Messzeitpunkt (MZIP 1) mit 0.079 signifikant von 0 (vgl. Tabelle 43). Die lernabhängige Veränderung des Deep-Structure Index (Erklärungsmodell) mit nur 0.005 ist nicht signifikant. Die Hypothese $H_{0[1,3]}$ wird beibehalten; die Ähnlichkeit der MODELLE zum Erklärungsmodell (Deep-

Tabelle 43: Lineares Wachstumsmodell der Deep-Structure, Erklärungsmodell (Studie 3)

Fixe Effekte	<i>Coefficient</i>	<i>SE</i>	<i>t-Wert</i>	<i>df</i>	<i>p</i>
Mean Initial Status π_{0i}	0.079	0.013	5.802	46	< .001
Mean Growth Rate π_{1i}	0.005	0.003	1.616	46	.113
Zufällige Effekte	<i>Variance Component</i>	<i>SD</i>	<i>df</i>	χ^2	<i>p</i>
Initial Status r_{0i}	0.00451	0.067	46	94.957	< .001
Growth Rate r_{1i}	0.00026	0.016	46	82.383	.001
Level-1 Error e_{ti}	0.00912	0.095			

Structure) verändert sich nicht während des Lernprozesses.

Die Ergebnisse für den Deep-Structure Index (Expertenmodell) zeigen ein ähnliches Bild auf. Zum ersten Messzeitpunkt (MZIP 1) liegt der Deep-Structure Index (Expertenmodell) mit 0.06 signifikant über 0. Wiederum kann jedoch keine signifikante lernabhängige Veränderung des Deep-Structure Index geschätzt werden (vgl. Tabelle 44). Folglich wird Hypothese $H_{0[1.3]}$ beibehalten; die Ähnlichkeit der MODELLE mit dem Expertenmodell (Deep-Structure) verändert sich während des Lernprozesses nicht.

Tabelle 44: Lineares Wachstumsmodell der Deep-Structure, Expertenmodell (Studie 3)

Fixe Effekte	Coefficient	SE	t-Wert	df	p
Mean Initial Status π_{0i}	0.061	0.010	5.863	46	< .001
Mean Growth Rate π_{1i}	0.002	0.002	0.796	46	.430

Zufällige Effekte	Variance Component	SD	df	χ^2	p
Initial Status r_{0i}	0.00337	0.05809	46	131.115	< .001
Growth Rate r_{1i}	0.00021	0.01456	46	115.471	< .001
Level-1 Error e_{ti}	0.00393	0.06268			

Auch in der dritten experimentellen Studie soll ein zweites Wachstumsmodell (*Level-2 Modell*) mögliche Einflüsse von zeitstabilen Drittvariablen aufklären (vgl. Raudenbush & Bryk, 2002). Hierfür wird das *Level-2* Submodell nach der allgemeinen Gleichung 8.1.3.2 konstruiert. Für jede der Ebenen der *SMD-Technologie* (Surface-, Matching- und Deep-Structure) wurde eine HLM-Analyse nach Gleichung 8.1.3.2 mit den Prädiktoren *GRUPPE* (0 = *Kontrollgruppe*, 1 = *self-guided learning*, 2 = *scaffolding-based learning*) und *VORWISSEN* (0 = *niedrigeres Vorwissen*, 1 = *höheres Vorwissen*) durchgeführt. Die Ergebnisse werden in Tabellenform in Anhang H dargestellt. Hinsichtlich der Surface-Structure und der experimentellen Variation *GRUPPE* ergibt sich mittels Substitution der Ergebnisse folgendes *Level-2* Submodell (Gleichung 8.3.3.2):

$$\begin{aligned}\hat{\pi}_{0i} &= 2.343 + 0.809GRUPPE_i \\ \hat{\pi}_{1i} &= 0.508 + 0.191GRUPPE_i\end{aligned}\quad [8.3.3.2]$$

Zum ersten Messzeitpunkt (MZIP 1) erreichen die Probanden der Bedingung *Kontrollgruppe* einen Surface-Structure Index von 2.343, welcher sich signi-

fikant von 0 unterscheidet. Der Surface-Structure Index der Probanden der Bedingung *self-guided learning* ist zum ersten Messzeitpunkt im Durchschnitt um 0.809 (3.152) höher. Probanden der Bedingung *scaffolding-based learning* erreichen einen durchschnittlichen Surface-Structure Index zum ersten Messzeitpunkt von 3.961 (2.343+0.809*2). Die Untersuchungsgruppen unterscheiden sich zum ersten Messzeitpunkt (MZP 1) jedoch nicht signifikant (vgl. Anhang H; 49,7 % aufgeklärte Varianz). Die lernabhängige Veränderung des Surface-Structure Index wird im zweiten Teil der Gleichung 8.3.3.2 angegeben. Für die lernabhängige Veränderung der Surface-Structure kann kein signifikanter Effekt berichtet werden. Hypothese $H_{0[2,1]}$ wird beibehalten; die lernabhängige Veränderung der Anzahl der Propositionen (Surface-Structure) der MODELLE unterscheidet sich nicht zwischen den Untersuchungsgruppen *Kontrollgruppe*, *self-guided learning* und *scaffolding-based learning*.

Die Ergebnisse bezüglich des Prädiktors *VORWISSEN* zeigen, dass sich die Probanden mit *niedrigerem Vorwissen* zum ersten Messzeitpunkt (MZP 1) signifikant von 0 unterscheiden. Der Unterschied zu Probanden mit *höherem Vorwissen* ist nicht signifikant (50,0 % aufgeklärte Varianz). Die Hypothese $H_{0[3,1]}$ wird beibehalten, welche besagt, dass das bereichsspezifische Vorwissen keinen Einfluss auf die lernabhängige Veränderung der Anzahl der Propositionen (Surface-Structure) hat (vgl. Anhang H).

Aus den Ergebnissen der HLM-Analyse für die Ebene der Matching-Structure ergibt sich durch Substitution folgende Gleichung (8.3.3.3) für den Prädiktor *GRUPPE*:

$$\begin{aligned}\hat{\pi}_{0i} &= 1.010 + 0.391GRUPPE_i \\ \hat{\pi}_{1i} &= 0.227 + 0.035GRUPPE_i\end{aligned}\quad [8.3.3.3]$$

Der detaillierten Ergebnisdarstellung in Anhang H ist zu entnehmen, dass sich die Veränderung des Durchmessers der MODELLE (Matching-Structure) zwischen den Probanden der Bedingungen *Kontrollgruppe*, *self-guided learning* und *scaffolding-based learning* nicht signifikant unterscheidet (26,9 % aufgeklärte Varianz). Folglich wird Hypothese $H_{0[2,2]}$ beibehalten; die lernabhängige Veränderung des Durchmesser (Matching-Structure) der MODELLE unterscheidet sich nicht zwischen den Untersuchungsgruppen.

Für den Prädiktor *VORWISSEN* konnte ein signifikanter Effekt für die Untersuchungsgruppe mit *niedrigerem Vorwissen* zum ersten Messzeitpunkt (MZP 1) und hinsichtlich deren lernabhängigen Veränderung festgestellt werden (vgl. Anhang H). Die Probanden mit *höherem Vorwissen* unterscheiden sich von den Probanden mit *niedrigerem Vorwissen* jedoch nicht bezüglich dem Wert zum ersten Messzeitpunkt und deren lernabhängigen Veränderung der Matching-Structure (22,9 % aufgeklärte Varianz). Somit wird die Hypothese $H_{0[3,2]}$ beibehalten. Das bereichsspezifische Vorwissen hat keinen Einfluss auf die lernabhängige Veränderung des Durchmesser (Matching-Structure).

$$\begin{aligned}\hat{\pi}_{0i} &= 0.061 + 0.013GRUPPE_i \\ \hat{\pi}_{1i} &= 0.001 + 0.003GRUPPE_i\end{aligned}\quad [8.3.3.4]$$

Gleichung 8.3.3.4 zeigt das *Level-2* Submodell nach einer weiteren HLM-Analyse mit der abhängigen Variable Deep-Structure (Erklärungsmodell). Die Ergebnisse in Anhang H zeigen keine signifikanten Unterschiede für die Untersuchungsgruppen (*Kontrollgruppe*, *self-guided learning* und *scaffolding-based learning*) bezüglich deren lernabhängigen Veränderung. Die Hypothese $H_{0[2,3]}$ wird beibehalten; die lernabhängige Veränderung der Ähnlichkeit der MODELLE mit dem Erklärungsmodell (Deep-Structure) unterscheidet sich nicht zwischen den Untersuchungsgruppen *Kontrollgruppe*, *self-guided learning* und *scaffolding-based learning*.

Die Ergebnisse für den Prädiktor *VORWISSEN* zeigen, dass sich die Probanden mit *niedrigerem* und *höherem Vorwissen* bezüglich deren lernabhängigen Veränderung nicht signifikant unterscheiden (vgl. Anhang H). Die Hypothese $H_{0[3,3]}$ wird beibehalten; das bereichsspezifische Vorwissen steht in keinem Zusammenhang mit der lernabhängigen Veränderung der Ähnlichkeit eines konstruierten MODELLS zum Erklärungsmodell (Deep-Structure).

Die folgende HLM-Analyse (Prädiktor *GRUPPE*) ergab für die Deep-Structure (Expertenmodell) in Gleichung 8.3.3.5 dargestelltes *Level-2* Submodell:

$$\begin{aligned}\hat{\pi}_{0i} &= 0.029 + 0.027GRUPPE_i \\ \hat{\pi}_{1i} &= 0.004 - 0.001GRUPPE_i\end{aligned}\quad [8.3.3.5]$$

Die Ergebnisdarstellung in Anhang H zeigt einen signifikanten Unterschied für die Untersuchungsgruppen (*Kontrollgruppe*, 0.029; *self-guided learning*, 0.056 und *scaffolding-based learning*, 0.083) bezüglich des Ausgangsniveau (MZIP 1). Hinsichtlich der lernabhängigen Veränderung unterscheiden sich die Probanden der einzelnen Untersuchungsgruppen jedoch nicht signifikant. Die Hypothese $H_{0[2,3]}$ wird beibehalten; die lernabhängige Veränderung der Ähnlichkeit der MODELLE mit dem Expertenmodell (Deep-Structure) unterscheidet sich nicht zwischen den Untersuchungsgruppen.

Für den Prädiktor *VORWISSEN* kann auch kein signifikanter Unterschied bezüglich der lernabhängigen Veränderung (Deep-Structure, Expertenmodell) zwischen Probanden mit *niedrigerem Vorwissen* (0.002) und Probanden mit *höherem Vorwissen* (0.0014) berichtet werden (vgl. Anhang H). Somit wird Hypothese $H_{0[3,3]}$ beibehalten: Das bereichsspezifische Vorwissen steht in keinem Zusammenhang mit der lernabhängigen Veränderung der Ähnlichkeit eines konstruierten MODELLS zum Expertenmodell (Deep-Structure).

8.4 ZUSAMMENFASSUNG DER ERGEBNISSE

Mittels dem *Oberflächen-Struktur-Vergleich* und der *SMD-Technologie* wurden in drei experimentellen Studien lernabhängige Veränderungen von MODELLEN diagnostiziert. Die 106 teilnehmenden Probanden konstruierten insgesamt 742 MODELLE, welche detailliert analysiert wurden. Zusammenfassend kann festgehalten werden, dass über die sieben Messzeitpunkte der drei experimentellen Studien lernabhängige Veränderungen der konstruierten MODELLE festgestellt werden konnten. Die deskriptive Betrachtung der Ergebnisse zeigt unterschiedliche Verläufe der lernabhängigen Veränderung der MODELLE hinsichtlich der experimentellen Bedingungen auf. Neben der lernabhängigen Veränderung der konstruierten MODELLE konnten Effekte von pädagogischen Interventionen während der Lernphase diagnostiziert werden. Im Gegensatz dazu konnte gezeigt werden, dass das deklarative Vorwissen der Probanden in allen drei experimentellen Studien keinen Einfluss auf die lernabhängige Veränderung der MODELLE hat.

Die Diskussion im folgenden Kapitel soll die umfangreichen statistischen Analysen und die detaillierte Ergebnisdarstellung dieses Kapitels aufgreifen.

Darüber hinaus werden die verwendeten Diagnoseverfahren *Oberflächen-Struktur-Vergleich* und *SMD-Technologie* hinsichtlich deren Praktikabilität kritisch beleuchtet.

9 DISKUSSION DER ERGEBNISSE

“Valid techniques for research on change have long perplexed behavioral scientists.” (Raudenbush & Bryk, 2002, p. 160)

Im Zentrum dieser Arbeit stand die *Diagnose lernabhängiger Veränderungen individueller Modellkonstruktionen*. Ausgehend von einer Einführung in die Theorie der mentalen Modelle (Seel, 1991), wurde der aktuelle Forschungsstand hinsichtlich geeigneter Verfahren zur Diagnose mentaler Modelle diskutiert. Die methodologische Analyse zeigt jedoch, dass mittels dieser Verfahren eine Diagnose lernabhängiger Veränderungen mentaler Modelle aus ökonomischen und versuchsplantechnischen Gründen oder hinsichtlich teils mangelnder Objektivität, Reliabilität und Validität nur eingeschränkt oder gar nicht durchführbar ist. Darauf aufbauend wurden die Verfahren *Oberflächen-Struktur-Vergleich* und *SMD-Technologie* speziell für eine Diagnose lernabhängiger Veränderungen mentaler Modelle entwickelt.

Die Diskussion der Ergebnisse fokussiert zunächst die empirische Überprüfung der Diagnoseverfahren *Oberflächen-Struktur-Vergleich* (Kapitel 9.1.1) und *SMD-Technologie* (Kapitel 9.1.2). In einem weiteren Schritt werden die Ergebnisse der drei experimentellen Studien analysiert und in Zusammenhang mit der Theorie der mentalen Modelle (Seel, 1991) gestellt (Kapitel 9.2). In Kapitel 9.3 folgt abschließend eine Synthese der empirischen Forschungsarbeit.

9.1 DIAGNOSEINSTRUMENTE

Grundvoraussetzung für eine solide empirische Forschung sind Instrumente, die die vorgegebenen Konstrukte objektiv, reliabel und valide messen (vgl. Seel, 1999a). In Kapitel 3 wurden die Instrumente *Verfahren des Lauten Denkens*, *Struktur-Lege-Techniken*, *Concept-Mapping-Tools* und *Test für Kausalmodelle* differenziert betrachtet und für die Fragestellung dieser Arbeit als nicht oder nur eingeschränkt geeignet bewertet. Dies gab Anlass für die Neuentwicklung der Diagnoseinstrumente *Oberflächen-Struktur-Vergleich* und *SMD-Technologie*. Erstes Ziel dieser Arbeit war es folglich, Gütekriterien für die neu entwickelten Diagnoseinstrumente zu bestimmen.

9.1.1 Oberflächen-Struktur-Vergleich

Der *Oberflächen-Struktur-Vergleich* (SSC: Surface-Structure-Comparison) eignet sich für die Diagnose augenscheinlicher Ähnlichkeiten graphisch externalisierter Modellkonstruktionen. Aus der Anwendung und empirischen Überprüfung des *Oberflächen-Struktur-Vergleichs* in den drei experimentellen Studien lassen sich (1) allgemeine Anforderungen und (2) grundsätzliche Einschränkungen aufzeigen.

(1) Grundvoraussetzung für die Anwendung des *Oberflächen-Struktur-Vergleichs* ist ein experimentelles Design mit einer für die Forschungsfrage adäquaten Anzahl von Messzeitpunkten. Weiter ist zu jedem Messzeitpunkt der aktuelle Lernstand (i.S. eines Messzeitpunktmodells) der Probanden in graphischer Form zu diagnostizieren. Dies erfolgte in den drei experimentellen Studien mittels *MS PowerPoint* (Microsoft Corporation, 2000)⁴⁴, was sowohl für die Schüler als auch für die Studierenden nach einer kurzen Schulung als unproblematisch einzustufen ist (vgl. Couné et al., 2003). Schließlich werden unabhängige *Model-Rater* für das Begutachtungsverfahren benötigt. In den drei empirischen Studien erwiesen sich zwischen 17 (dritte experimentelle Studie) und 31 (erste experimentelle Studie) *Model-Rater* als ausreichend.

Die Test-Retest-Reliabilität des *Oberflächen-Struktur-Vergleichs* konnte in der ersten und zweiten experimentellen Studie geprüft werden. In beiden Studien konnte eine hohe Zuverlässigkeit des *Oberflächen-Struktur-Vergleichs* nachgewiesen werden (vgl. Kapitel 7). Die *Model-Rater* hatten keine Mühe, die MODELLE hinsichtlich deren Ähnlichkeiten einzuschätzen. Außerdem erwies sich das Verfahren sowohl für die *Model-Rater* als auch für den empirischen Forscher als sehr ökonomisch und wenig zeitaufwendig (vgl. Ifenthaler & Seel, 2005).

(2) Grundlegende Einschränkungen ergeben sich bezüglich des Informationsgehalts der gewonnenen Daten. So ist es ohne weiteres möglich, die durchschnittliche Veränderungswahrscheinlichkeit zwischen einzelnen Messzeitpunkten zu berechnen und somit die Progression der Modellkonstruktionen während des Lernprozesses zu veranschaulichen (vgl. Kapitel 8). Einschrän-

⁴⁴ Es sei an dieser Stelle darauf hingewiesen, dass eine Diagnose auch mit *Paper-and-Pencil-Tests* oder *Concept-Mapping-Tools* erfolgen kann.

kend ist jedoch anzumerken, dass mittels *Oberflächen-Struktur-Vergleich* ausschließlich die oberflächlichen Veränderungen der Modellkonstruktionen diagnostiziert werden. Die Analyse bietet folglich keine Informationen bezüglich der Struktur, Komplexität und semantischen Qualität der Modellkonstruktion.

Die Anwendung des *Oberflächen-Struktur-Vergleichs* empfiehlt sich daher nur, wenn (1) eine schnelle, reliable und ökonomische Analyse umfangreicher Daten gefordert wird und (2) die Einschränkungen im Informationsgewinn mit der Forschungsfrage vereinbar sind.

9.1.2 SMD-Technologie

Im Gegensatz zum *Oberflächen-Struktur-Vergleich* ermöglicht die *SMD-Technologie* eine umfangreiche und vielschichtige Diagnose lernabhängiger Veränderungen individueller Modellkonstruktionen. Für die Entwicklung der automatisierten *SMD-Technologie* (vgl. Kapitel 4) wurden (1) theoretische, (2) methodologische und (3) technologische Aspekte berücksichtigt.

Aus (1) theoretischer Sicht wurde die *SMD-Technologie* im Hinblick auf das Konstrukt der mentalen Modelle (Seel, 1991) entwickelt. In Kapitel 2 und 3 konnte gezeigt werden, dass nicht ein mentales Modell, sondern nur dessen externalisierte Repräsentation diagnostiziert werden kann. Die Externalisierung mentaler Modelle wird im Rahmen dieser Arbeit als bewusster Prozess einer Person in Form einer Codierung und Kommunikation mittels geeigneter Zeichen- und Symbolsysteme verstanden. Personen zeigen jedoch Präferenzen hinsichtlich der Verwendung von Zeichen- und Symbolsystemen bzw. unterscheiden sich hinsichtlich ihrer semiotischen Kompetenz (vgl. Seel, 1999b). Die *SMD-Technologie* wird diesen Anforderungen gerecht, da sowohl graphische als auch textbasierte Forschungsdaten in die automatisierte Analyse einfließen können (vgl. Kapitel 4). Für die Transformation der grafischen Repräsentationen in die für die *SMD-Technologie* notwendige Listenform (vgl. Kapitel 4), konnte für die drei experimentellen Studien dieser Arbeit eine hohe Interrater-Reliabilität nachgewiesen werden (Kapitel 7.3.5).

Mittels der drei Ebenen Surface-, Matching- und Deep-Structure, können die konstruierten Modelle detailliert hinsichtlich deren *relationalen*, *strukturellen* und *semantisch-strukturellen Eigenschaften* diagnostiziert werden. Die für die menta-

le Modellbildung (Seel, 1991) notwendigen Prozesse der Konstruktion, Modifikation und Reorganisation individueller Modellkonstruktionen, wird durch die drei Ebenen der *SMD-Technologie* präzise erfasst und durch drei eindeutig interpretierbare Indizes dokumentiert (vgl. Kapitel 4).

(2) Bei der Entwicklung der *SMD-Technologie* wurde zudem die Annahme berücksichtigt, dass sich mentale Modelle während des Lernprozesses von Novizenmodellen hin zu Erklärungs- oder Expertenmodellen entwickeln (vgl. Snow, 1990 und Kapitel 2). Die Diagnose der sukzessiven Modellvervollständigung (Seel, 1991) bedingt, dass die *SMD-Technologie* Daten mehrerer Messzeitpunkte erfassen und analysieren kann. Erschwerend kommt aus methodologischer Sicht der Umgang mit Veränderungsdaten hinzu (vgl. Kapitel 5). Wie die historische Synopse der Veränderungsmessung (Kapitel 5.2) aufzeigen konnte, ergeben sich bei Fragestellungen mit Veränderungsdaten methodologische Probleme, die aus unterschiedlicher Sicht gelöst werden können. Für die Datenanalyse der drei experimentellen Studien dieser Arbeit konnte die **Hierarchisch Lineare Modellierung** (Raudenbush & Bryk, 2002) als aussichtsreiches statistisches Verfahren verwendet werden (vgl. Kapitel 5, 7 und 8).

Aus (3) technologischer Sicht wurde die automatisierte Analyse der *SMD-Technologie* plattformunabhängig als webbasierte Anwendung implementiert. Es wird daher keine Installation einer zusätzlichen Software durch den Diagnostiker erfordert, was die Usability dieser Software entscheidend positiv beeinflusst (vgl. Tergan, 2004). Als weiterer Vorteil erwiesen sich die umfangreichen Funktionen der MySQL-Datenbank. Neben der Archivierung großer Datenmengen, können für die *SMD-Technologie* beliebig viele Erklärungs- und Expertenmodelle zur Verfügung gestellt werden, die eine umfassende Diagnose individueller Modellkonstruktionen ermöglichen.

Die zeitliche Dauer der Analysephase ist von der Komplexität des Referenzmodells und der Versuchspersonenmodelle, sowie von der Anzahl der einzelnen Modelle abhängig. Die Analyse aller Daten der drei experimentellen Studien der vorliegenden Arbeit erfolgte in zirka drei Minuten. Dem Diagnostiker werden alle Ergebnisse der *SMD-Technologie* (Usernamen, VPN-Nummer, MZP-Nummer, Surface-, Matching-, Deep-Structure, Auswertungszeit) in der SQL-Datenbank zur Verfügung gestellt. Durch die Exportfunktion können die Daten für weitere Analysen mit Statistiksoftware aufbereitet werden. Zusätzlich

wird eine dynamische HTML-Seite mit allen Ergebnissen inklusive einer Repräsentation der individuellen MODELLE generiert und automatisch dargestellt. Jede Rerepräsentation der individuellen MODELLE wird auf dem Webserver der *SMD-Technologie* als Bilddatei gespeichert und steht für weitere diagnostische Zwecke zur Verfügung.

Aus der Überprüfung der Gütekriterien (Kapitel 7.4) kann auf eine hohe Objektivität, Reliabilität und Validität der *SMD-Technologie* geschlossen werden. Die empirischen Analysen zur Überprüfung der *Validität* der *SMD-Technologie* zeigen sowohl *konvergente* als auch *divergente Validität* bezüglich der verwendeten Außenkriterien. Außerdem konnte eine *Kreuzvalidierung* mit mehreren quantitativen und qualitativen Diagnoseinstrumenten zur Erfassung von externalisierten Modellen erfolgen (vgl. Kapitel 7.3.5).

Insgesamt kann festgehalten werden, dass die *SMD-Technologie* als ein zuverlässiges und ökonomisches Verfahren zur Erfassung und Analyse individueller Modellkonstruktionen eingestuft werden kann. Die umfassenden Annahmen der Theorie der mentalen Modelle (Seel, 1991) erforderte die umfangreiche Berücksichtigung von theoretischen, methodologischen und technologischen Aspekten bei der Entwicklung und Implementation der *SMD-Technologie*. Die folgende Diskussion der Ergebnisse der drei experimentellen Studien verdeutlichen, dass die individuellen Modellkonstruktionen mit weniger komplexen Diagnoseverfahren nur bruchstückhaft zu analysieren wären.

9.2 ERGEBNISSE DER EXPERIMENTELLEN STUDIEN

Wie bereits mehrfach verdeutlicht wurde, konnten im Rahmen der drei experimentellen Studien nicht mentale Modelle, sondern nur externalisierte MODELLE der Probanden diagnostiziert werden. Hierdurch ergeben sich im Hinblick auf die Theorie der mentalen Modelle (Seel, 1991) Einschränkungen bezüglich der Interpretation der vorliegenden Untersuchungsergebnisse. Zum einen mussten die Probanden vorgegebene Zeichen- und Symbolsysteme für die Kommunikation bzw. Externalisierung ihrer mentalen Modelle verwenden, um eine Vergleichbarkeit der Daten sicherzustellen. Trotz einer umfassenden Einführung in die Verwendung der Zeichen- und Symbolsysteme muss davon ausgegangen werden, dass diese Form der Kommunikation von mentalen Mo-

dellen nicht für alle Probanden geeignet war (Seel, 1999b). Zum anderen ist es wahrscheinlich, dass während der Lernphase die mentalen Modelle weit umfangreicher als die externalisierten Modelle waren. Die qualitative Reduktion der MODELLE kann sowohl auf der relationalen, strukturellen, aber auch auf der semantischen Ebene stattfinden. Da die pädagogisch-psychologische Forschung zu mentalen Modellen jedoch keine anderen Forschungszugänge zur Verfügung stellt, sollen die Ergebnisse der drei experimentellen Studien vor dem Hintergrund der gerade ausgeführten Einschränkungen diskutiert werden. Dabei sollen die Fragestellungen und Hypothesen (Kapitel 6) mit den Ergebnissen (Kapitel 8) und den theoretischen Ausführungen (Kapitel 2 bis 5) verknüpft werden.

Die multimediale, entdecken-lassende Lernumgebung zum Thema *Geologie* (erste experimentelle Studie) und *Geophysik* (zweite und dritte experimentelle Studie) stellte den Probanden eine Vielzahl von Informationen in unterschiedlichen Repräsentationsformaten zur Verfügung (vgl. Kapitel 7.1). Die Navigation innerhalb der Lernumgebung wurde den Probanden frei gestellt, es konnte jedoch gezeigt werden, dass der *Modellbankkasten* (MoBuKi) signifikant häufiger als die anderen Bereiche der multimedialen Lernumgebung benutzt wurde (vgl. Dummer & Ifenthaler, 2005). Die komplexe Problemstellung der Lernumgebung war nach der Theorie der mentalen Modelle (Seel, 1991) nur dann zu lösen, wenn Probanden mit Hilfe der zur Verfügung stehenden Informationen und ihrem bereits vorhandenen Weltwissen subjektiv plausible mentale Modelle konstruieren.

9.2.1 Lernabhängige Veränderung der Modelle

Die lernabhängige Veränderung mentaler Modelle vollzieht sich nach Johnson-Laird (1989) von *Novizenmodellen* hin zu *Expertenmodellen*. Folglich wurden während der Lernphase der drei experimentellen Studien Unterschiede zwischen den einzelnen Messzeitpunkten erwartet.

Die deskriptiven Ergebnisse des *Oberflächen-Struktur-Vergleichs* weisen auf eine lernabhängige Veränderung der mentalen Modelle hin. In allen drei experimentellen Studien konnte eine hohe Veränderungswahrscheinlichkeit zu Beginn der Lernphase (zwischen dem ersten und dritten Messzeitpunkt) festgestellt werden. Dies kann mehrere Ursachen haben. Zum einen kann davon

ausgegangen werden, dass die Probanden mit Hilfe der bereitgestellten Informationen der multimedialen Lernumgebung neue, subjektiv plausiblere mentale Modelle konstruiert haben. Andererseits kann die anfängliche Veränderung der mentalen Modelle auch damit zusammenhängen, dass die Probanden erst allmählich den Anforderungen der Kommunikation ihrer mentalen Modelle mittels der geforderten Zeichen- und Symbolsysteme gerecht wurden.

Nach zwei Stunden Lernphase konnte in allen drei experimentellen Studien mittels *Oberflächen-Struktur-Vergleich* nur noch eine geringe Veränderungswahrscheinlichkeit nachgewiesen werden. Die immer kleiner werdende Veränderungswahrscheinlichkeit der mentalen Modelle zum Ende des Lernprozesses deutet auf eine Schematisierung der Modelle hin (vgl. Seel, 1991, Hanke, 2006). Da dieser Effekt in allen drei experimentellen Studien zu etwa der gleichen Zeit während der Lernphase eingetreten ist, können folgende Schlussfolgerungen gezogen werden. Ad hoc Konstruktionen, wie die der mentalen Modelle, werden durch mehrfache Konstruktion zu identischen, komplexen Problemstellungen schematisiert (Gick & Holyoak, 1983). Hierzu bleibt anzumerken, dass im Gegensatz zu mentalen Modellen ein Schema im Langzeitgedächtnis gespeichert und somit immer wieder abgerufen werden kann (vgl. Kapitel 2.1.1).

Auf Grund des eingeschränkten Informationsgewinns des *Oberflächen-Struktur-Vergleichs* (Kapitel 9.1) liefern die Ergebnisse keine Aussagen über die qualitativen Veränderungen der mentalen Modelle. Aussichtsreicher sind hier die Ergebnisse der *SMD-Technologie*. Die präzise Diagnose der MODELLE mittels der Surface-, Matching- und Deep-Structure zu jedem Messzeitpunkt ermöglicht eine detaillierte Betrachtung der lernabhängigen Veränderungen.

Wie die Ergebnisse der Surface-Structure zeigen, nimmt die Summe der Propositionen der mentalen Modelle während der Lernphase in allen drei experimentellen Studien signifikant zu. Neue Informationen werden folglich bei subjektiver Plausibilität in ein erneut konstruiertes mentales Modell integriert. Wie bereits ausgeführt ist zu vermuten, dass dieser Prozess erst dann abflacht, wenn keine neuen, subjektiv plausible Informationen zur Verfügung stehen, oder eine Schematisierung des Modells eingetreten ist.

Die Ergebnisse der HLM-Analysen der Matching-Structure verdeutlichen, dass sich auch die Komplexität der mentalen Modelle in allen drei experimen-

tellen Studien signifikant verändert hat. Durch die Integration neuer Informationen werden die Zusammenhänge in einem mentalen Modell während der Lernphase komplexer, was mit dem Durchmesser der mentalen Modelle abgebildet werden kann. Der Zusammenhang der Matching-Structure mit der Surface-Structure liegt auf der Hand. Werden bei einer erneuten Konstruktion eines mentalen Modells neue Informationen verarbeitet, so nimmt nicht nur die Anzahl der Verwendeten Propositionen zu, sondern auch die Komplexität der mentalen Modelle.

Letztlich stellt sich die Frage, ob sich die mentalen Modelle auch hinsichtlich deren semantischen Qualität in Bezug auf ein *Erklärungs-* und/oder *Expertenmodell* während des Lernprozesses verändern. Für die Deep-Structure konnte kein einheitliches Bild einer lernabhängigen Veränderung für die drei experimentellen Studien gefunden werden. In der ersten experimentellen Studie werden die mentalen Modelle dem *Erklärungsmodell* während der Lernphase signifikant ähnlicher. Die Ähnlichkeit zum *Expertenmodell* veränderte sich während der Lernphase nicht signifikant. In der zweiten experimentellen Studie konnte eine signifikante Veränderung der Ähnlichkeit zum *Expertenmodell*, aber nicht zum *Erklärungsmodell* festgestellt werden. Für die dritte experimentelle Studie kann keine signifikante Veränderung berichtet werden. Aus diesen Ergebnissen können folgende Annahmen abgeleitet werden. (1) Die bereitgestellten Informationen in der multimedialen Lernumgebung waren nicht ausreichend, um eine semantische Ähnlichkeit mit dem *Erklärungs-* bzw. *Expertenmodell* zu erzeugen. (2) In der Lernumgebung wurden Lernhilfen nicht ausreichend angeboten, um evt. Misskonzeptionen der Probanden zu vermeiden (vgl. Snow, 1990; Seel et al., 2000). Die komplette Lernzeit (3) war zu kurz, um die relevanten Information für eine umfassende Lösung der komplexen Problemstellung zu erarbeiten. (4) Die komplexe Problemstellung war für die Zielgruppe nicht geeignet, da notwendige Komponenten (Informationskomponente, Reizkomponente, Reaktionskomponente, Rückmeldungs-komponente) bei der Gestaltung der Lernaufgabe nicht ausreichend berücksichtigt wurden (vgl. Seel, 1981).

Zusammenfassend kann festgehalten werden, dass während des Lernprozesses eine lernabhängige Veränderung der mentalen Modelle diagnostiziert werden konnte. Diese beschränkt sich in den drei experimentellen Studien jedoch hauptsächlich auf die Summe der verwendeten Propositionen (Surface-Structure) und auf die Komplexität der mentalen Modelle (Matching-

Structure). Die semantische Analyse der mentalen Modelle (Deep-Structure) zeigte, dass die inhaltliche Qualität der mentalen Modelle nur geringfügig zunimmt. In allen drei Studien waren die Probanden nicht in der Lage, eine adäquate Lösung der komplexen Problemstellung abzuliefern. Die Ergebnisse stimmen mit der Annahme von Larkin (1983) überein, wonach Novizenmodelle oberflächliche Eigenschaften der Inhaltsdomäne darstellen und Expertenmodelle tiefer greifende Eigenschaften berücksichtigen. Anschließend ist daher zu klären, ob die pädagogischen Interventionen der drei experimentellen Studien den Lernprozess effektiv beeinflussen konnten.

9.2.2 Pädagogische Interventionen

Eine Vielzahl empirischer Arbeiten haben sich damit beschäftigt, ob es grundsätzlich möglich ist, durch instruktionale Maßnahmen auf die Veränderung mentaler Modelle einzuwirken (vgl. Kapitel 2.1.2, Anzai & Yokoyama, 1984; Mayer, 1989; Seel & Dinter, 1995; Al-Diban, 2002). Das experimentelle Design dieser Arbeiten ermöglicht jedoch in den wenigsten Fällen eine präzise Diagnose der Effekte pädagogischer Interventionen über mehrere Messzeitpunkte. An diese Kritik anknüpfend wurde das Design der drei experimentellen Studien dieser Arbeit mit sieben Messzeitpunkten und unterschiedlichen Untersuchungsbedingungen (*scaffolding-based learning*, *self-guided learning* und *Kontrollgruppe*) verwirklicht (vgl. Kapitel 7).

Die Ergebnisse der HLM-Analysen (Level-1: *Zeitfaktor bzw. Messzeitpunkte*; Level-2: *Variation der Untersuchungsbedingung*) zeigen, dass sich die Untersuchungsgruppen bezüglich deren lernabhängigen Veränderung nur in der zweiten experimentellen Studie auf den Ebenen der Surface- und Matching-Structure signifikant unterscheiden (vgl. Kapitel 8 bzw. Anhang G). In diesen beiden Fällen erreichen die Probanden der Untersuchungsbedingung *scaffolding-based learning* signifikant höhere Surface- und Matching-Structure Indizes, als die Probanden der Untersuchungsbedingung *self-guided learning*. Alle übrigen HLM-Analysen zeigten keine signifikanten Effekte bezüglich deren lernabhängigen Veränderung auf. Gründe hierfür liegen vermutlich im didaktischen Design der multimedialen Lernumgebung. Weder das individuelle Feedback, noch die (automatisierten) prozessorientierten Lernhilfen in der multimedialen Lernumgebung (vgl. Kapitel 7) zeigten in den drei experimentellen Studien Wirkung. Es

kann schließlich festgehalten werden, dass die in den drei experimentellen Studien realisierten pädagogischen Interventionen nicht zum richtigen Zeitpunkt platziert waren und/oder nicht die bestmögliche Form hatten. Da zu diesem Zeitpunkt keine empirischen Studien über pädagogische Interventionen in entdecken-lassenden *modellzentrierten* Lernumgebungen bekannt sind, besteht in diesem Feld aus Sicht des Autors ein erhöhter Forschungsbedarf.

9.2.3 Effekte von Lernervoraussetzungen

Eine weitere Forschungsfrage im Rahmen der empirischen Untersuchung geht von der Annahme Ausubel's (1968, p. vi) aus: „The most important single factor influencing learning is what the learner already knows“. In Bezug auf die Verarbeitung von neuen Informationen in komplexen Problemlösesituationen wird angenommen, dass das individuelle Vorwissen einer Person in engem Zusammenhang mit der Konstruktion, Modifikation und Reorganisation mentaler Modelle steht (vgl. Seel, 1991). Im Rahmen der drei experimentellen Studien wurde zu Beginn das deklarative Wissen der relevanten Inhaltsdomäne mittels eines Multiple-Choice-Fragebogens erhoben. Mittels HLM (Level-1: *Zeitfaktor bzw. Messzeitpunkte*; Level-2: *individuelles Vorwissen*) sollten mögliche Einflüsse des *deklarativen Vorwissens* auf die lernabhängige Veränderung mentaler Modelle analysiert werden.

Die Ergebnisse in Kapitel 8 bzw. Anhang F bis H konnten die Annahmen nicht bestätigen. Ein einziger signifikanter Effekt kann für die Veränderung der Deep-Structure (Expertenmodell) der zweiten experimentellen Studie berichtet werden. Die Probanden mit höherem *deklarativem Vorwissen* veränderten ihre mentalen Modelle in Bezug auf die Ähnlichkeit zum Expertenmodell stärker als die Probanden mit niedrigerem deklarativen Vorwissen. Folglich kann der Einfluss von *deklarativem Vorwissen* auf die lernabhängige Veränderung mentaler Modelle nicht ganz ausgeschlossen werden. Gerade in komplexen Problemlöseaufgaben kann ein ausgeprägtes Vorwissen die mentale Modellbildung entscheidend beeinflussen. In weiteren Forschungsarbeiten bleibt dennoch zu überprüfen, ob neben deklarativem Vorwissen auch prozedurale und semiotische Kompetenzen die mentale Modellbildung entscheidend beeinflussen.

9.3 SYNTHESE DER EMPIRISCHEN FORSCHUNGSARBEIT

Ziel der empirischen Forschungsarbeit war die *Diagnose lernabhängiger Veränderung mentaler Modelle* in einer problemorientierten, multimedialen, entdecken-lassenden Lernumgebung. Diese sollte den Probanden „Gelegenheiten zum Nachdenken“ (Stolurow, 1973) bieten. Für die Diagnose der MODELLE wurden die Instrumente *Oberflächen-Struktur-Vergleich* und *SMD-Technologie* entwickelt. In drei experimentellen Studien mit Längsschnittdesign wurde die Progression mentaler Modelle, der Einfluss von pädagogischen Interventionen und Effekte von Lernervoraussetzungen untersucht. Zentrale Ergebnisse der drei experimentellen Studien sind:

1. Der *Oberflächen-Struktur-Vergleich* erwies sich als schnelles, ökonomisches und reliables Verfahren für die Diagnose augenscheinlicher Ähnlichkeiten graphisch externalisierter Modellkonstruktionen. Der eingeschränkte Informationsgehalt der gewonnen Daten muss jedoch mit der empirischen Forschungsfrage vereinbar sein.
2. Die *SMD-Technologie* wurde im Hinblick auf die Theorie der mentalen Modelle, auf die methodologischen Anforderungen der Veränderungsmessung und auf die aktuellen technologischen Realisierungsmöglichkeiten entwickelt.
3. Individuelle Modellkonstruktionen konnten durch die drei Ebenen der *SMD-Technologie* präzise analysiert werden. Die unterschiedlichen Perspektiven der Surface-, Matching- und Deep-Structure ermöglichen eine umfassende Diagnose lernabhängiger Veränderungen mentaler Modelle.
4. Die Überprüfung der Gütekriterien ergab für die *SMD-Technologie* eine hohe Objektivität, Reliabilität und Validität. Darüber hinaus kann die automatisierte *SMD-Technologie* als schnelles, leicht handhabbares und sehr ökonomisches Verfahren eingestuft werden.
5. Die *Hierarchisch Lineare Modellierung* erwies sich als aussichtsreiches statistisches Verfahren zur Modellierung individueller Veränderungen. Für Veränderungsmessungen wird das Mehrebenen-Regressionsmodell aus einer ersten Ebene (Level-1) mit dem Zeitfaktor und aus einer zweiten Ebene (Level-2) mit den zu untersuchenden Subjekten gebildet.

6. Die *lernabhängige Veränderung mentaler Modelle* konnte in den drei experimentellen Studien bezüglich der Summe der verwendeten Propositionen und der Komplexität der Modelle festgestellt werden. Insgesamt waren die mentalen Modelle dem *Erklärungsmodell* ähnlicher als dem *Expertenmodell*.
7. Die *pädagogischen Interventionen* der Lernumgebung zeigten nur kurzfristige oder gar keine Wirkung.
8. Der Einfluss des *deklarativen Vorwissens* der Probanden auf die individuelle Modellkonstruktion konnte in den drei experimentellen Studien nicht eindeutig nachgewiesen werden.
9. Die empirischen Ergebnisse konnten in den drei experimentellen Studien mit unterschiedlichen Stichproben und Inhaltsdomänen repliziert werden.

10 ZUSAMMENFASSUNG UND AUSBLICK

„Only by measuring individual change it is possible to document each person's progress and, consequently, to evaluate the effectiveness of instruction.” (Seel, 1999a, p. 180)

Im Zentrum der vorliegenden Arbeit stand die Frage, ob das entwickelte Verfahren der *SMD-Technologie* geeignet ist, lernabhängige Veränderungen individueller Modellkonstruktionen auf der relationalen, strukturellen und semantischen Ebene zu diagnostizieren.

Die *Theorie der mentalen Modelle* (Seel, 1991) geht davon aus, dass sich individuelle Modellkonstruktionen in Abhängigkeit eines Lernprozesses verändern. Diese lernabhängige Veränderung mentaler Modelle vollzieht sich nach Johnson-Laird (1989) von Novizenmodellen hin zu Expertenmodellen. Mit der Einführung eines *Erklärungsmodells*, welches die minimal notwendigen Informationen für eine korrekte Problemlösung enthält, soll die Annahme einer lernabhängigen Veränderung mentaler Modelle präzisiert werden. Ein im Gegensatz zum Expertenmodell weniger komplexes bzw. abstraktes *Erklärungsmodell* scheint in einer komplexen Problemlösesituation schneller und effektiver zu sein, da die Komplexität eines Expertenmodells den Problemlöseprozess behindern kann.

Die *pädagogische Diagnose mentaler Modelle*, bzw. die Frage nach reliablen und validen Instrumenten stellt ein zentrales Problem der Forschung zu mentalen Modellen dar. Da mentale Modelle nicht unmittelbar diagnostiziert werden können, müssen diese von Personen mittels geeigneter Zeichen- und Symbolsysteme kommuniziert werden. Dieser Prozess wird als Externalisierung verstanden. Mehrere Instrumente konnten einer kritischen Analyse hinsichtlich deren Eignung für eine pädagogische Diagnose mentaler Modelle unterzogen werden. Die teilweise sehr zeitaufwendigen Verfahren weisen keine oder nur geringe Gütekriterien auf.

Die *SMD-Technologie* wurde speziell für die Diagnose lernabhängiger Veränderung mentaler Modelle entwickelt. Individuelle Modellkonstruktionen werden durch die *SMD-Technologie* auf drei Ebenen diagnostiziert. Diese erfassen die *relationale* (Surface-Structure), *strukturelle* (Matching-Structure) und *se-*

mantisch-strukturelle (Deep-Structure) Dimension der MODELLE. Die automatisierte *SMD-Technologie* erwies sich in der Durchführung als praktikables, leicht handhabbares und ökonomisches Verfahren. Außerdem konnte mittels empirischen Analysen für die *SMD-Technologie* eine hohe Objektivität, Reliabilität und Validität bestimmt werden.

Insbesondere die Lehr-Lern-Forschung befasst sich mit dem sukzessiven Aufbau und der Veränderung psychischer Dispositionen bei der Informationsverarbeitung. Um präzise Aussagen über diese komplexen Zusammenhänge treffen zu können, führt aus methodologischer Sicht an einer *Veränderungsmessung* kein Weg vorbei. Die speziellen Probleme der *Veränderungsmessung* können jedoch Forschungsergebnisse verfälschen und schließlich zu unbrauchbaren Interpretationen führen. Diese Dilemmata können jedoch mittels geeigneter Instrumente, auf die Forschungsfrage abgestimmter experimenteller Designs und umfangreicher testtheoretischer Modelle umgangen werden. Für die Fragestellungen dieser Arbeit erwies sich die **H**ierarchisch **L**ineare **M**odellierung als tragfähiger Ansatz.

Im Rahmen dieser Arbeit wurden *drei experimentelle Studien* mit unterschiedlichen Inhaltsdomänen durchgeführt. Insgesamt nahmen an den Studien $N = 106$ Probanden (Schüler und Studierende) teil. Die modular aufgebaute, multimediale Lernumgebung sollte den Probanden „Gelegenheiten zum Nachdenken“ bieten. Aufgabe der Probanden war es, eine komplexe Problemstellung zu lösen. Die Lösung sollte in Form von MODELLEN (i.S. benannter Knoten und Relationen) kommuniziert werden. Während der Lernphase wurden diese individuellen Modellkonstruktionen an sieben Messzeitpunkten erhoben. Die 742 MODELLE dienten als Datengrundlage für die umfangreichen empirischen Analysen mittels der Instrumente *Oberflächen-Struktur-Vergleich* und *SMD-Technologie*.

Die *Ergebnisse* zeigten, dass mittels der *SMD-Technologie* eine präzise Diagnose lernabhängiger Veränderungen mentaler Modelle auf den drei Ebenen Surface-, Matching- und Deep-Structure möglich ist. Die lernabhängige Veränderung mentaler Modelle konnte in den drei experimentellen Studien auf der *relationalen* und *strukturellen Ebene* festgestellt werden. Außerdem wurde eine größere *semantisch-strukturelle Ähnlichkeit* der MODELLE zum *Erklärungsmodell* als zum *Expertenmodell* diagnostiziert. Die pädagogischen Interventionen der multimedialen Lernumgebung zeigten nur kurzfristige oder gar keine Wirkung.

Auch das deklarative Vorwissen der Probanden hatte keinen Einfluss auf die Konstruktion der mentalen Modelle. Die empirischen Ergebnisse der drei experimentellen Studien konnten mit unterschiedlichen Stichproben und in unterschiedlichen Inhaltsdomänen repliziert werden.

Die bisherige Anwendung der *SMD-Technologie* beschränkte sich auf die drei experimentellen Laboruntersuchungen dieser Arbeit. Aus dieser Einschränkung ergeben sich an dieser Stelle *offene Fragen* für weitere Forschungsaktivitäten hinsichtlich der *SMD-Technologie* und der *Diagnose lernabhängiger Veränderungen mentaler Modelle*.

1. Ist der Einsatz der *SMD-Technologie* in längsschnittlichen Untersuchungen im Bereich von Unterricht und Weiterbildungsveranstaltungen möglich?
2. Eignet sich jegliche Art von *Inhaltsdomäne* für eine umfassende Analyse mittels der *SMD-Technologie*?
3. Können die Unterschiede zwischen einem *Erklärungsmodell* und einem *Expertenmodell* in verschiedenen Inhaltsdomänen klar definiert werden?
4. Wie muss eine *pädagogische Intervention* in einer modellzentrierten Lernumgebung konstruiert sein, damit diese effektiv die individuelle Modellkonstruktion unterstützt?
5. Tritt bei länger angelegten Lernprozessen eine Schematisierung ein und bleibt diese bei Stabilitätsuntersuchungen konstant?

Für *weitere Forschungsaktivitäten* mit der *SMD-Technologie* erscheint es an dieser Stelle sinnvoll, dass das Verfahren in einigen Bereichen modifiziert wird. Die *Weiterentwicklung der SMD-Technologie* sollte neben dem bereits *automatisierten Diagnosemodul* ein *Konstruktionsmodul* enthalten. Dieses *Konstruktionsmodul* ermöglicht zum einen die Konstruktion, Modifikation und Reorganisation von MODELLEN in Form von benannten Knoten und Relationen. Darüber hinaus soll aber auch die Möglichkeit bestehen, dass Anwender ihre mentalen Modelle in Form von natürlicher Sprache (Fließtext) kommunizieren. In Kombination mit dem *Diagnosemodul* können Anwender der *SMD-Technologie* unmittelbar nach der Konstruktion eines MODELLS zur Problemstellung eine umfangreiche Rückmeldung bezüglich eines Erklärungs- und / oder Expertenmodells erhalten. Die Forschungs- und Anwendungsbereiche dieser erweiterten *SMD-Technologie* ließen sich somit um ein vielfaches erweitern.

LITERATUR

- Al-Diban, S. (2002). *Diagnose mentaler Modelle*. Hamburg: Kovac.
- American Psychological Association (2005). *Publication Manual of the American Psychological Association*. Washington, DC: American Psychological Association.
- Anderson, R.C. (1984). Some reflections on the acquisition of knowledge. *Educational Researcher*, 13(9), 5-10.
- Anzai, Y., Yokoyama, T. (1984). Internal models in physics problem solving. *Cognition and Instruction*, 1, 397-450.
- Arbeitsgruppe empirisch-pädagogische Forschung (2005). *Tagung der Sektion Empirische Bildungsforschung. Veränderungsmessung und Längsschnittstudien. Neue Datengrundlagen für Bildungsforschung und Bildungsmonitoring*. Berlin: Tagungsband.
- AT&T (2006). Graphviz - Graph Visualization Software [Computer Software]. Retrieved May 25, 2005, from <http://www.graphviz.org>
- Atkinson, J.W. (1964). *An introduction to motivation*. New York, NY: Van Nostrand.
- Ausubel, D.P. (1968). *Educational psychology. A cognitive view*. New York, NY: Holt, Rinehart, Winston.
- Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (2000). *Multivariate Analysemethoden: eine anwendungsorientierte Einführung*. Berlin: Springer.
- Bandura, A. (1979). *Sozial-kognitive Lerntheorie*. Stuttgart: Klett.
- Bartholomew, D.J. (1967). *Stochastic models for social processes*. New York, NY: Wiley.
- Bartlett, F.C. (1932). *Remembering. A study in experimental and social psychology*. Cambridge, UK: Cambridge University Press.

- Bentler, P.M. (1986). EQS - Ein Ansatz zur Analyse von Strukturgleichungsmodellen für normal- bzw. nichtnormal verteilte quantitative Variablen. In C. Moebus & W. Schneider (Eds.), *Strukturmodelle für Längsschnittdaten und Zeitreihen. LISREL, Pfad- und Varianzanalyse* (pp. 27-56). Bern: Huber.
- Bentler, P.M. (1992). *EQS: Structural equations program manual* (Vers. 3.0). Los Angeles, CA: BMDP.
- Bereiter, C. (1963). Some persisting dilemmas in the measurement of change. In C.W. Harris (Ed.), *Problems in measuring change* (pp. 3-20). Madison, WI: The University of Wisconsin Press.
- Bierhoff, H. W. (2005). *Sozialpsychologie. Ein Lehrbuch*. Stuttgart: Kohlhammer.
- Bock, R.D. (1963). Multivariate analysis of variance id repeated measurements. In C.W. Harris (Ed.), *Problems in measuring change* (pp. 85-103). Madison, WI: The University of Wisconsin Press.
- Bonato, M. (1990). *Wissensstrukturierung mittels Struktur-Ge-Techniken. Eine graphentheoretische Analyse von Wissensnetzen*. Frankfurt am Main: Lang.
- Bortz, J. (1993). *Statistik für Sozialwissenschaftler*. Berlin: Springer.
- Bortz, J. & Döring, N. (1995). *Forschungsmethoden und Evaluation*. Berlin: Springer.
- Bracht, G.H. & Glass, G.V. (1975). Die externe Validität von Experimenten. In R. Schwarzer & K. Steinhagen (Eds.), *Adaptiver Unterricht. Zur Wechselwirkung von Schülermerkmalen und Unterrichtsmethoden* (pp. 64-93). München: Kösel-Verlag.
- Brosius, F. (2002). *SPSS 11*. Bonn: MITP-Verlag.
- Brown, A.L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *Journal of the Learning Sciences*, 2(2), 141-178.
- Bruner, J.S. (1964). The course of cognitive growth. *American Psychologist*, 19, 1-16.

- Bryk, A.S. & Raudenbush, S.W. (1987). Application of hierarchical linear models to assessing change. *Psychological Bulletin*, 101, 147-158.
- Bryk, A.S. & Raudenbush, S.W. (1992). *Hierarchical linear models: Applications and data analysis methods*. Newbury Park, CA: Sage Publications.
- Bryk, A.S., Raudenbush, S.W. & Congdon, R.T.J. (1994). *Hierarchical Linear Modeling with HLM/2L and HLM/3L programs*. Chicago, IL: Scientific Software International.
- Bühner, M. (2004). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson Studium.
- Campbell, D.T. & Stanley, J.C. (1966). *Experimental and quasi-experimental designs for research*. Chicago, IL: Rand McNally.
- Coleman, E.B. (1998). Using explanatory knowledge during collaborative problem solving in science. *Journal of the Learning Sciences*, 7(3-4), 387-427.
- Collins, L. & Horn, J. (1991). *Best methods for analysis of change: Recent advances, unanswered questions, future directions*. Washington, DC: American Psychological Association.
- Collins, L.M. & Sayer, A.G. (2001). *New methods for the analysis of change*. Washington, DC: American Psychological Association.
- Couné, B. (2006, in Vorbereitung). *Motivationale Bedingungsfaktoren von Lernleistung. Untersuchung von Erwartungs- und Wertkompetenzen in einer modellbasierten Lernumgebung*. Freiburg: Universitäts-Dissertation.
- Couné, B., Hanke, U., Ifenthaler, D. & Seel, N.M. (2003). *Modellkonstruktionen beim Problemlösen im Kontext entdeckenden Lernens: Eine Studie zur Implementierung theoretisch-begründeter Instruktionsprinzipien. Erster Bericht aus dem Forschungsprojekt „Modell-begründetes Lernen und Lehren. Multimediale Lernumgebungen als Gelegenheiten zum Nachdenken*. Freiburg: Institut für Erziehungswissenschaft.
- Couné, B., Hanke, U., Ifenthaler, D. & Seel, N.M. (2004). *Modellkonstruktionen beim Problemlösen im Kontext entdeckenden Lernens: Eine Studie zur Implementie-*

- rung theoretisch-begründeter Instruktionsprinzipien. Zweiter Bericht aus dem Forschungsprojekt „Modell-begründetes Lernen und Lehren. Multimediale Lernumgebungen als Gelegenheiten zum Nachdenken.* Freiburg: Institut für Erziehungswissenschaft.
- Craik, K. (1943). *The nature of explanation.* Cambridge, UK: Cambridge University Press.
- Cronbach, L.J. & Furby L. (1970). How should we measure “change” – or should we? *Psychological Bulletin*, 74(1), 68-80.
- Dietrich, R. (1977). *Psychodiagnostik: Grundlagen und Probleme.* München: Reinhardt.
- Dinter, F.R. (1993). *Mentale Modelle als Konstrukt der empirischen Erziehungswissenschaft.* Saarbrücken: Universität Dissertation.
- Ditton, H. (1998). *Mehrebenenanalyse. Grundlagen und Anwendungen des Hierarchisch Linearen Modells.* Weinheim: Juventa.
- Doll, J. & Knopf, H. (2001). Studentische Beurteilung der Lehrveranstaltungen ihres berufsbegleitenden Studienganges "Psychologie - Lehramt an Gymnasien" in Abhängigkeit von Berufserfahrung und Selbstwirksamkeit. In G. Krampen & H. Zayer (Eds.), *Psychologiedidaktik und Evaluation III: Konzepte, empirische Befunde und Erfahrungsberichte zur psychologischen Aus-, Fort- und Weiterbildung* (pp. 263-281). Bonn: Deutscher Psychologen Verlag.
- Doll, J. & Prenzel, M. (2001). Das DFG-Schwerpunktprogramm "Bildungsqualität von Schule" (BIQUA): Bedingungen mathematischer, naturwissenschaftlicher und Überfachlicher Kompetenzen. In TIMSS: *Impulse für Schule und Unterricht - Forschungsbefunde, Reforminitiativen, Praxisberichte und Videodokumente* (pp. 99-104). Bonn: BMBF.
- Doll, J. & Prenzel, M. (2003). Themenschwerpunkte und Ergebnisse des DFG-Schwerpunktprogramms "Bildungsqualität von Schule". In I. Gogolin & R. Tippelt (Eds.) *"Innovation durch Bildung", Beiträge zum 18. Kongress der Deutschen Gesellschaft für Erziehungswissenschaft* (pp. 225-240). Opladen: Leske + Budrich.

- Dörner, D. (1974). *Die kognitive Organisation beim Problemlösen*. Bern: Huber.
- Dummer, P. & Ifenthaler, D. (2005). Planning and assessing navigation in model-centered learning environments. Why learners often do not follow the path laid out for them. In G. Chiazese, M. Allegra, A. Chifari & S. Ottaviano (Eds.), *Methods and technologies for learning* (pp. 327-334). Southampton: WIT Press.
- Dutke, S. (1994). *Mentale Modelle: Konstrukte des Wissens und Verstehens*. Göttingen: Verlag für Angewandte Psychologie.
- Eckert, A. (1998). *Kognition und Wissensdiagnose. Die Entwicklung und empirische Überprüfung des computerunterstützten wissensdiagnostischen Instrumentariums Netzwerk-Elaborierungs-Technik (NET)*. Lengerich: Pabst.
- Eckert, A. (2000). Die Netzwerk-Elaborierungs-Technik (NET) – Ein computerunterstütztes Verfahren zur Diagnose komplexer Wissensstrukturen. In H. Mandl & F. Fischer (Eds.), *Wissen sichtbar machen – Wissensmanagement mit Mapping-Techniken* (pp. 138-157). Göttingen: Hogrefe.
- Eckert, A. & Niegemann, H. M. (1995). Computerunterstützte Diagnose von Wissensstrukturen und deren Veränderung. In R. Arbinger & R.S. Jäger (Eds.), *Zukunftsperspektiven empirisch-pädagogischer Forschung* (pp. 352-362). Landau: Verlag Empirische Pädagogik.
- Eckert, T. & Tippelt, R. (2001). Wandel des Angebots in der Erwachsenenbildung. Regionalstudie in Freiburg. In E. Nuissl & E. Schultz (Eds.), *Systemevaluation und Politikberatung* (pp. 164-182). Bielefeld: Bertelsmann.
- Eid, M. (2003). Veränderungsmessung und Kausalanalysen. In M. Jerusalem & H. Weber (Eds.), *Psychologische Gesundheitsförderung – Diagnostik und Prävention* (pp. 105-120). Göttingen: Hogrefe.
- Eigler, G. (1990). Funktionaler Analphabetismus: Einführung. *Unterrichtswissenschaft*, 18(2), 98-100.
- Engel, W. & Reinecke, J. (1996). *Analysis of change: Advanced techniques in panel data analysis*. Berlin: de Gruyter.

- Fahrenberg, J. (1968). Aufgaben und Methoden der psychologischen Verlaufsanalyse (Zeitreihenanalyse). In K.J. Groffmann & K.H. Wewetzer (Eds.), *Person als Prozeß* (pp. 41-82). Stuttgart: Huber.
- Fahrenberg, J., Kuhn, M., Kulick, B. & Myartek, M. (1977). Methodenentwicklung für psychologische Zeitreihenstudien. *Diagnostica*, 23, 15-36.
- Fischer, G.H. & Formann, A.K. (1982). Veränderungsmessung mittels linear-logistischer Modelle. *Zeitschrift für Differenzielle und Diagnostische Psychologie*, 3(2), 75-99.
- Funke, J. (1985). Steuerung dynamischer Systeme durch Aufbau und Anwendung subjektiver Kausalmodelle. *Zeitschrift für Psychologie*, 193(4), 443-465.
- Funke, J. (1990). Systemmerkmale als Determinanten des Umgangs mit dynamischen Systemen. *Sprache & Kognition*, 9(3), 143-153.
- Funke, J. (1991). Solving complex problems: Exploration and control of complex problems. In R.J. Sternberg & P.A. Frensch (Eds.), *Complex problem solving: Principles and mechanisms* (pp. 185-222). Hillsdale, NJ: Lawrence Erlbaum.
- Funke, J. (1999). Komplexes Problemlösen: Ein Blick zurück und nach vorne. *Psychologische Rundschau*, 50, 194-197.
- Funke, J. (2003). *Problemlösendes Denken*. Stuttgart: Kohlhammer.
- Gage, N.L. & Berliner, D.C. (1996). *Pädagogische Psychologie*. Weinheim: Psychologie Verlags Union.
- Gaito, J. & Wiley, D.E. (1963). Univariate analysis of variance procedures in the measurement of change. In C.W. Harris (Ed.), *Problems in measuring change* (pp. 60-84). Madison, WI: The University of Wisconsin Press.
- Gardner, H. (1985). Die Entwicklung von Symbolkompetenz bei Kindern. *Zeitschrift für Semiotik*, 7, 319-333.
- Gentner, D. & Stevens, A.L. (1983). *Mental models*. Hillsdale, NJ: Lawrence Erlbaum.

- Gick, M., Holyoak, K.J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, 15, 1-38.
- Greeno, J.G. (1989). Situations, mental models, and generative knowledge. In D. Klahr & K. Kotovsky (Eds.), *Complex information processing* (pp. 285-318). Hillsdale, NJ: Lawrence Erlbaum.
- GRADAP (1990). *Graph definition and analysis package user's manual*. Groningen: IEC-ProGamma.
- Goldstein, H. (1979). *The design and analysis of longitudinal studies. Their role in the measurement of change*. London: Academic Press.
- Hanke, U. (2006). *Externale Modellbildung als Hilfe bei der Informationsverarbeitung und beim Lernen*. Freiburg: Universitäts-Dissertation.
- Harary, F. (1974). *Graphentheorie*. München: Oldenbourg.
- Harris, C.W. (1963). *Problems in measuring change*. Madison, WI: The University of Wisconsin Press.
- Heckhausen, H. (1963). *Hoffnung und Furcht in der Leistungsmotivation*. Meisenheim am Glan: Hain.
- Heckhausen, H. (1989). *Motivation und Handeln. Lehrbuch der Motivationspsychologie*. Berlin: Springer.
- Held, C., Kanuff, M. & Vosgerau, G. (2006). *Mental models and the mind. Current development in cognitive psychology, neuroscience, and philosophy of mind*. Amsterdam: Elsevier.
- Helmreich, R. (1977). *Strategien zur Auswertung von Längsschnittdaten. Ein Beitrag zur Messung von Veränderung in der empirischen Sozialforschung*. Stuttgart: Klett.
- Hofmann, D.A., Jacobs, R. & Baratta, J.E. (1993). Dynamic criteria and the measurement of change. *Journal of Applied Psychology*, 2, 194-204.
- Holtzmann, W.H. (1963). Statistical models for the study of change in the single case. In C.W. Harris (Ed.), *Problems in measuring change* (pp. 199-211). Madison, WI: The University of Wisconsin Press.

- Holyoak, K.J. (1984). Mental models in problem solving. In J.R. Anderson & S. Kosslyn (Eds.), *Tutorials in learning and memory: essays in honor of Gordon Bower* (pp. 193-218). San Francisco, CA: W.H. Freeman and Company.
- Horst, P. (1963). Multivariate models for evaluating change. In C.W. Harris (Ed.), *Problems in measuring change* (pp. 104-121). Madison, WI: The University of Wisconsin Press.
- Hox, J. (2002). *Multilevel analysis. Techniques and applications*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Ifenthaler, D. & Seel, N.M. (2005). The measurement of change: Learning-dependent progression of mental models. *Technology, Instruction, Cognition and Learning*, 2(4), 317-336.
- Ingenkamp, K. (1997). *Lehrbuch der pädagogischen Diagnostik*. Weinheim: Beltz.
- Jäger, A.O., Süss, H. & Beauducel, A. (1997). *Berliner Intelligenzstruktur Test*. Göttingen: Hogrefe.
- Johnson, T.E., O'Connor, D.L. Spector, J.M., Ifenthaler, D. & Pirnay-Dummer, P. (2006). *Comparative study of mental model research methods: Relationships among ACSMM, SMD, MITOCAR & DEEP methodologies*. Paper submitted to the CMC 2006 - Second International Conference on Concept Mapping, September 5-8, San José, Costa Rica.
- Johnson-Laird, P.N. (1983). *Mental models. Towards a cognitive science of language, inference, and consciousness*. Cambridge, UK: Cambridge University Press.
- Johnson-Laird, P.N. (1989). Mental Models. In M.I. Posner (Ed.), *Foundations of Cognitive Science* (pp. 469-499). Cambridge, MA: MIT Press.
- Jonassen, D.H. (1987). Assessing cognitive structure: Verifying a method using pattern notes. *Journal of Research and Development in Education*, 20(3), 1-14.
- Jöreskog, K.G. & Sörbom, D. (1989). *LISREL 7. A guide to the program and applications*. Chicago, IL: SPSS Inc.
- Jöreskog, K.G., Sörbom, D., du Toit, S. & du Toit, M. (1999). *LISREL 8: New statistical features*. Chicago, IL: Scientific Software International Inc.

- Kerlinger, F.N. (1986). *Foundation of behavioral research*. Fort Worth: Harcourt, Brace, Jovanovich.
- Klauer, K.J. (1978). *Handbuch der Pädagogischen Diagnostik. Band 1*. Düsseldorf: Pädagogischer Verlag Schwann.
- Klauer, K.J. (2005). *Das Experiment in der pädagogisch-psychologischen Forschung. Eine Einführung*. Münster: Waxmann.
- Kleber, E.W. (1992). *Diagnostik in pädagogischen Handlungsfeldern: Einführung in Bewertung, Beurteilung, Diagnose und Evaluation*. München: Juventa-Verlag.
- Kleiter, E.F. (1987). Stand und Perspektive der Veränderungsmessung aus formal-gegenständlicher Sicht. *Zeitschrift für Pädagogische Psychologie*, 1(4), 235-258.
- Kluwe, R.H. (1988). Methoden der Psychologie zur Gewinnung von Daten über menschliches Wissen. In H. Mandl & H. Spada (Eds.), *Wissenspsychologie* (pp. 359-385). München: Psychologie Verlags Union.
- Krauth, J. (1981). Statistische Methoden zur Veränderungsmessung. In U. Baumann, H. Berbalk & G. Seidenstücker (Eds.), *Klinische Psychologie: Trends in Forschung und Praxis 4* (pp. 98-131). Bern: Huber.
- Krems, J. (1995). Cognitive flexibility and complex problem solving. In P.A. Frensch & J. Funke (Eds.), *Complex problem solving. The European perspective* (pp. 201-218). Hillsdale, NJ: Lawrence Erlbaum.
- Langer, W. (2004). *Mehrebenenanalyse. Eine Einführung für Forschung und Praxis*. Wiesbaden: VS Verlag.
- Langheine, R. & Rost, J. (1996). Latent-Class Analyse. In E. Erdfelder, R. Mausfeld, T. Meisinger & G. Rudinger (Eds.), *Handbuch quantitative Methoden* (pp. 315-332). Weinheim: Beltz.
- Larkin, J.H. (1983). The role of problem representation in physics. In D. Gentner & A.L. Stevens (Eds.), *Mental models* (pp. 75-98). Hillsdale, NJ: Lawrence Erlbaum.

- Le Ny, J.-F. (1993). Wie kann man mentale Repräsentationen repräsentieren? In: J. Engelkamp & T. Pechmann (Eds.), *Mentale Repräsentation* (pp. 31-39). Bern: Huber.
- Lewin, K. (1942). Feldtheorie des Lernens. In C.F. Graumann (Ed.), *Kurt-Lewin-Werkausgabe. Band 4: Feldtheorie* (pp. 157-185). Bern: Huber.
- Lienert, G.A. & Raatz, U. (1994). *Testaufbau und Testanalyse*. Weinheim: Beltz.
- Lin, D. (1998). *An information-theoretic definition of similarity*. International Conference on Machine Learning, Madison, WI.
- Lord, F.M. (1963). Elementary models for measuring change. In C.W. Harris (Ed.), *Problems in measuring change* (pp. 21-38). Madison, WI: The University of Wisconsin Press.
- Mandl, H. & Fischer, F. (2000). Mapping-Techniken und Begriffsnetze in Lern- und Kooperationsprozessen. In H. Mandl & F. Fischer (Eds.), *Wissen sichtbar machen – Wissensmanagement mit Mapping-Techniken* (pp. 3-12). Göttingen: Hogrefe.
- Mandl, H., Friedrich, H.F. & Horn, A. (1988). Theoretische Ansätze zum Wissenserwerb. In H. Mandl & H. Spada (Eds.), *Wissenspsychologie* (pp. 123-160). München: Psychologie Verlags Union.
- MannheimResearchCompany (2005). *Mannheimer Netzwerk Elaborations Technik - MaNET*. Retrieved March 23, 2005, from <http://www.marescom.net>
- Mayer, R.E. (1989). Models for understanding. *Review of Educational Research*, 59(1), 43-64.
- McCloskey, M. (1983). Naïve theories of motion. In D. Gentner & A.L. Stevens (Eds.), *Mental models*, (pp. 299-324). Hillsdale, NJ: Lawrence Erlbaum.
- Microsoft Corporation (2000). *Microsoft Office XP. PowerPoint 2000* [Computer Software]. Redmond, WA. Microsoft Corporation.

- Minsky (1975). A framework for representing knowledge. In P.H. Winston (Ed.), *The psychology of computer vision*, (pp. 211-277). New York; NY: McGraw-Hill.
- Moskowitz, D.S. & Hershberger, S.L. (2002). *Modeling intraindividual variability with repeated measures data. Methods and applications*. Mahwah, NJ: Lawrence Erlbaum.
- Muthen, B.O. (1987). *LISCOMP: Analysis of linear structural equations with a comprehensive measurement model*. Mooresville, IN: Scientific Software, Inc.
- Nenninger, P. (1980). Anwendungsmöglichkeiten der Graphentheorie in der Erziehungswissenschaft. *Zeitschrift für Empirische Pädagogik*, 4, 8-106.
- Norman, D.A. (1983). Some observations on mental models. In D. Gentner & A.L. Stevens (Eds.), *Mental models* (pp. 7-14). Hillsdale, NJ: Lawrence Erlbaum.
- Novak, J.D. (1998). *Learning, creating, and using knowledge: concept maps as facilitative tools in schools and corporations*. Mahwah, NJ: Lawrence Erlbaum.
- O'Connor, D.L., Johnson, T.E. (2004). *Measuring team cognition: Concept mapping elicitation as a means of constructing team shared mental models in an applied setting*. First International Conference on Concept Mapping, September 14-17, 2004, Pamplona, Spain.
- Oehr, A. (1889). Experimentelle Studien zur Individualpsychologie. In E. Kraepelin (Ed.), *Psychologische Arbeiten*, 1 (pp. 92-151).
- Pawlik, K. (1976). Modell- und Praxisdimensionen psychologischer Diagnostik. In K. Pawlik (Ed.), *Diagnose der Diagnostik: Beiträge zur Diskussion der Psychologischen Diagnostik in der Verhaltensmodifikation* (pp. 13-43). Stuttgart: Klett.
- Pellegrino, J.W., Chudowsky, N. & Glaser, R. (2001). *Knowing what students know: the science and design of educational assessment*. Washington, DC: National Academy Press.

- Penner, D.E. (2001). Cognition, computers and synthetic science: Building knowledge and meaning through modelling. *Review of Research in Education*, 25, 1-35.
- Perrez, M. (2005). *Lehrbuch Klinische Psychologie - Psychotherapie*. Bern: Huber.
- Petermann, F. (1978). *Veränderungsmessung*. Stuttgart: Kohlhammer.
- Petermann, F. (1986). Probleme und neuere Entwicklungen der Veränderungsmessung – ein Überblick. *Diagnostica*, 32(1), 4-16.
- Petermann, F., Hehl, F.-J. & Schneider, W. (1977). Veränderungsmessung im Rahmen der klassischen Testtheorie. In F. Petermann (Ed.), *Methodische Grundlagen klinischer Psychologie* (pp. 129-154). Weinheim: Beltz.
- Piaget, J. (1950). *La construction du réel chez l'enfant*. Neuchâtel, Suisse: Delachaux et Niestlé S.A.
- Piaget, J. (1975). *Nachahmung, Spiel und Traum*. Stuttgart: Klett.
- Piaget, J. (1976). *Die Äquilibration der kognitiven Strukturen*. Stuttgart: Klett.
- Pirnay-Dummer, P. (2006, in Vorbereitung). *Expertise und Modellbildung: MITO-CAR*. Freiburg: Universitäts-Dissertation.
- Pirnay-Dummer, P., Ifenthaler, D. & Seel, N.M. (2006). *Role of learning strategies in mental model building processes*. San Francisco: Paper presented at the AREA 2006 (55.081).
- Putz-Osterloh, W., Bott, B. & Köster, K. (1990). Models of learning in problem solving – are they transferable to tutorial systems? *Computers in Human Behavior*, 6, 83-96.
- Raudenbush, S.W. & Bryk, A.S. (2002). *Hierarchical linear models. Applications and data analysis methods*. Thousand Oaks, CA: Sage Publications.
- Raudenbush, S.W., Bryk, A.S., Cheong, Y.F., Congdon, R. & du Toit, M. (2004). *HLM 6: Hierarchical linear & nonlinear modeling*. Lincolnwood, IL: Scientific Software International, Inc.

- Reiska, P. (2005). *Experimente und Computersimulationen im naturwissenschaftlichen Unterricht: Ein empirischer Vergleich*. Frankfurt am Main: Lang.
- Renkl, A. & Gruber, H. (1995). Erfassung von Veränderung: Wie und wieso? *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 27(2), 173-190.
- Renn, H. (1973). *Die Messung von Sozialisationswirkungen*. München: Oldenbourg.
- Rips, L.J. (1987). Mental muddles. In M. Brand, & R.M. Harnish (Eds.), *The representation of knowledge and belief* (pp. 259-286). Tuscon, AZ: University of Arizona Press.
- Ross, B.H. (1987). This is like that: The use of earlier problems and the separation of similarity effects. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 14(4), 629-639.
- Ross, B.H. (1989). Distinguishing types of superficial similarities. Different effects on the access and use of earlier problems. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 15, 456-468.
- Rost, J. (2004). *Lehrbuch Testtheorie – Testkonstruktion*. Bern: Huber.
- Rost, J. (2005). *Interpretation und Bewertung pädagogisch-psychologischer Studien*. Weinheim: Beltz.
- Rothgang, G.-W. (2003). *Entwicklungspsychologie*. Stuttgart: Kohlhammer.
- Rouse, W.B. & Morris, N.M. (1986). On looking into the black box: Prospects and limits in the search for mental models. *Psychological Bulletin*, 100(3), 349-363.
- Rumelhart, D.E. & Ortony, A. (1977). The representation of knowledge in memory. In R.C. Anderson, J.R. Spiro, & W.E. Montague (Eds.), *Schooling and the acquisition of knowledge* (pp. 99-135). Hillsdale, NJ: Lawrence Erlbaum.
- Rumelhart, D.E. & McClelland, J.L. (1986). *Parallel distributed processing. Explorations in the microstructure of cognition. Vol. 1: Foundations*. Cambridge, MA: The MIT Press.

- Rumelhart, D.E. & Norman, D.A. (1988). Representation in memory. In R.C. Atkinson, R.J. Herrnstein, G. Lindzey & R.D. Luce (Eds.), *Steven's handbook of experimental psychology* (pp. 99-135). Hillsdale, NJ: Lawrence Erlbaum.
- Scandura, J.M. (1988). Role of relativistic knowledge in intelligent tutoring. *Computers in Human Behavior*, 5, 53.64.
- Schank, R.C. (1982). *Dynamic memory. A theory of reminding and learning in computers and people*. Cambridge, UK: Cambridge University Press.
- Schank, R.C. (1999). *Dynamic Memory revisited*. Cambridge, UK: Cambridge University Press.
- Schank, R.C. & Abelson, R. (1977). *Scripts, plans, goals and understanding*. Hillsdale, NJ: Lawrence Erlbaum.
- Scheele, B. & Groeben, N. (1984). *Die Heidelberger Struktur-Lege-Technik (SLT). Eine Dialog-Konsens-Methode zur Erhebung Subjektiver Theorien mittlerer Reichweite*. Weinheim: Beltz.
- Scheele, B. & Groeben, N. (1988). *Dialog-Konsens-Methoden zur Rekonstruktion Subjektiver Theorien: Die Heidelberger Struktur-Lege-Technik (SLT), konsensuale Ziel-Mittel-Argumentation und kommunikative Flussdiagramm-Beschreibung von Handlungen*. Tübingen: Francke Verlag.
- Schenk, K. (2003). *Effekte metakognitiver Trainings auf Lernen und Problemlösen*. Dresden: Sachsoprint.
- Seashore, C.E. & Kent, G.H. (1905). Periodicity and progressive change in continuous mental work. *Psychological Review Monogr. Suppl.* 6, 47-101.
- Seel, N.M. (1981). *Lernaufgaben und Lernprozesse*. Stuttgart: Kohlhammer.
- Seel, N.M. (1991). *Weltwissen und mentale Modelle*. Göttingen: Hogrefe.
- Seel, N.M. (1997). Pädagogische Diagnose mentaler Modelle. In H. Gruber & A. Renkl (Eds.), *Wege zum Können: Determinanten des Kompetenzerwerbs* (pp. 116-137). Bern: Huber.

- Seel, N.M. (1999a). Educational diagnosis of mental models: Assessment problems and technology-based solutions. *Journal of Structural Learning and Intelligent Systems*, 14(2), 153-185.
- Seel, N.M. (1999b). Educational semiotics: School learning reconsidered. . *Journal of Structural Learning and Intelligent Systems*, 14(1), 11-28.
- Seel, N.M. (2001). *Modell-begründetes Lernen und Leben. Multimediale Lernumgebungen als „Gelegenheiten zum Nachdenken“: Nenantrag an die Deutsche Forschungsgemeinschaft*. Freiburg: Institut für Erziehungswissenschaft.
- Seel, N.M. (2003a). *Psychologie des Lernens. Lehrbuch für Pädagogen und Psychologen*. München: Reinhardt.
- Seel, N.M. (2003b). Model-centered learning and instruction. *Technology, Instruction, Cognition and Learning*, 1(1), 59-85.
- Seel, N.M. (2003c). Editorial Board Compendium II: Goals, Achievements, and Research Issues. *Technology, Instruction, Cognition and Learning*, 1(3), 242-251.
- Seel, N.M., Dinter, F.R. (1995). Instruction and mental model progression: Learner-dependent effects of teaching strategies on knowledge acquisition and analogical transfer. *Educational Research and Evaluation*, 1(1), 4-35.
- Seel, N.M., Al-Diban, S. & Blumschein, P. (2000). Mental models and instructional planing. In J.M. Spector & T.M. Anderson (Eds.), *Integrated and holistic perspectives on learning, instruction and technology: Understanding complexity* (pp. 129-158). Dordrecht: Kluwer Academic Publishers.
- Seel, N.M., Ifenthaler, D. & Pirnay-Dummer, P. (2006, in Vorbereitung). Mental models and problem solving: Technological solutions for the measurement and assessment of the development of expertise. In J. Strobel, D. Jonassen & W. Hung (Eds.), *Learning by modelling systems*. Mahwah, NJ: Lawrence Erlbaum.

- Simon, H.A. & Kaplan, C.A. (1989). Foundations of cognitive science. In M.I. Posner (Ed.), *Foundations of Cognitive Science* (pp. 1-47). Cambridge, MA: MIT Press.
- Singer, J.D. & Willett, J.B. (2003). *Applied longitudinal data analysis. Modeling change and event occurrence*. Oxford, NY: Oxford University Press.
- Smith, J.P., di Sessa, A.A. & Roschelle, J. (1993). Miconceptions reconceived: A constructivist analysis of knowledge in transition. *The Journal of the Learning Sciences*, 3(2), 115-163.
- Snow, R.E. (1990). New approaches to cognitive and conative assessment in education. *International Journal of Educational Research*, 14(5), 455-473.
- Spada, H. (1983). Die Analyse von Veränderungen im Rahmen unterschiedlicher testtheoretischer Modelle. In W.-R. Minsal & R. Scheller (Eds.), *Brennpunkte der Klinischen Psychologie* (pp. 83-105). München: Kösel-Verlag.
- Spector, J.M. (2006). Introduction to the special issue on models, simulations and learning in complex domains. *Technology, Instruction, Cognition and Learning*, 3(3-4), 199-203.
- Spector, J.M., Dennen, V.P. & Koszalka, T.A. (2005). Individual and collaborative construction of causal concept maps: An online technique for learning and assessment. In G. Chiazese, M. Allegra, A. Chifari & S. Ottaviano (Eds.), *Methods and technologies for learning* (pp. 223-227). Southampton: WIT Press.
- Spector, J.M., Dennen, V.P. & Koszalka, T.A. (2006). Causal maps, mental models and assessing acquisition of expertise. *Technology, Instruction, Cognition and Learning*, 3(1-2), 167-183.
- Spiro, R.J., Coulston, R.L., Feltovich, P.J. & Anderson, D.K. (1988). *Cognitive flexibility theory. Advanced knowledge acquisition in ill-structured domains. Technical report 441*. Champaign, IL: University of Illinois, Center for the Study of Reading.

- SPSS (2006). Statistical package for the social sciences. SPSS für Windows, Version 14 [Computer Software]. Chicago, IL: SPSS, Inc.
- Stachowiak, H. (1973). *Allgemeine Modelltheorie*. Berlin: Springer.
- Stelzl, I. (2005). *Fehler und Fallen der Statistik. Für Psychologen, Pädagogen und Sozialwissenschaftler*. Münster: Waxmann.
- Stieglitz, R.-D. (1986). *Erfassung von Veränderungen. Theoretische und empirische Beiträge*. Berlin: Oberhofer.
- Stolurow, L.M. (1973). Lernumwelten oder Gelegenheiten zum Nachdenken. In W. Edelstein & D. Hopf (Eds.). *Bedingungen des Bildungsprozesses: psychologische und pädagogische Forschungen zum Lehren und Lernen in der Schule* (pp. 351-398). Stuttgart: Klett.
- Stracke, I. (2004). *Einsatz computerbasierter Concept Maps zur Wissensdiagnose in der Chemie. Empirische Untersuchungen am Beispiel des Chemischen Gleichgewichts*. Münster: Waxmann.
- Tack, W.H. (1980). Zur Theorie psychometrischer Verfahren. Formalisierung der Erfassung von Situationsabhängigkeit und Veränderung. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 1(2), 87-106.
- Tent, L. & Stelzl, I. (1993). *Pädagogisch-psychologische Diagnostik. Band 1. Theoretische und methodische Grundlagen*. Göttingen: Hogrefe.
- Tergan, S.-O. (2004). *Was macht E-Learning erfolgreich? Grundlagen und Instrumente der Qualitätsbeurteilung*. Berlin: Springer.
- Tittmann, P. (2003). *Graphentheorie. Eine anwendungsorientierte Einführung*. München: Carl Hanser Verlag.
- Tversky, A. (1977). Features of similarity. *Psychological Review*, 84(4), 327-352.
- van de Pol, F. Langheine, R. & de Jong, W. (1991). *PANNMARK user manual. PANel analysis using MARKov chains*. Voorburg: Central Bureau of Statistics.
- van Someren, M.W., Barnard, Y.F. & Sandberg, J.A. (1994). *The think aloud method: a practical guide to modelling cognitive processes*. London: Academic Press.

- von Eye, A. (1990a). *Statistical methods in longitudinal research. Volume I. Principles and structuring change*. San Diego, CA: Academic Press.
- von Eye, A. (1990b). *Statistical methods in longitudinal research. Volume II. Time series and categorical longitudinal data*. San Diego, CA: Academic Press.
- von Sengenbusch, P. (2002). Modell des Energie- und Materialflusses in einem Ökosystem. Retrieved February 24, 2003, from <http://www.biologie.uni-hamburg.de/b-online/d54/54b.htm#06a>
- Vögele, E. (2004). *Aufgabenverständnis und Verlauf kognitiver Lernstrategienutzung : eine empirische Studie an der Hochschule*. Freiburg: Universitäts-Dissertation.
- Weber, S. (1994). *Vorwissen in der betriebswirtschaftlichen Ausbildung. Eine struktur- und inhaltsanalytische Studie*. Wiesbaden: Deutscher Universitäts-Verlag.
- Weber, S. & Schumann, M. (2000). Das Concept Mapping Software Tool (COMASOTO) zur Diagnose strukturellen Wissens. In H. Mandl & F. Fischer (Eds.), *Wissen sichtbar machen – Wissensmanagement mit Mapping-Techniken* (pp. 158-198). Göttingen: Hogrefe.
- Weiner, B. (1976). *Theorien der Motivation*. Stuttgart: Klett.
- Wild, K.-P. (2000). *Lernstrategien im Studium. Strukturen und Bedingungen*. Münster: Waxman.
- Wild, K.-P., Schiefele, U. (1994). Lernstrategien im Studium. Ergebnisse zur Faktorenstruktur und Reliabilität eines neuen Fragebogens. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 15, 185-200.
- Wild, K.-P., Schiefele, U., & Winteler, A. (1992). *LIST. Ein Verfahren zur Erfassung von Lernstrategien im Studium*. Neubiberg: Universität der Bundeswehr, Institut für Erziehungswissenschaft und Pädagogische Psychologie (Gelbe Reihe: Arbeiten zur Empirischen Pädagogik und Pädagogischen Psychologie, Nr.. 20).
- Willett, J.B. (1988). Questions and answers in the measurement of change. *Review of Research in Education*, 15, 345-422.

-
- Willeit, J.B. (1989). Some results on reliability for the longitudinal measurement of change: Implications for the design of studies of individual growth, *Educational and Psychological Measurement*, 49, 587-602.
- Wilson, J.R., Rutherford, A. (1989). Mental models: Theory and application in human factors. *Human Factors*, 31, 617-634.

ABBILDUNGSVERZEICHNIS

Abbildung 1: Prinzip der Äquilibration bezogen auf Schemata und mentale Modelle	10
Abbildung 2: Das Verhältnis von mentalen, konzeptuellen und instruktionalen Modellen in Bezug auf die Objekt- und Ereigniswelt (Seel, 2003a, p. 261)	14
Abbildung 3: Lernabhängige Veränderung mentaler Modelle nach Johnson-Laird (1989) und Snow (1990)	18
Abbildung 4: Erweiterte Darstellung der lernabhängigen Veränderung mentaler Modelle	21
Abbildung 5: Prozess der Externalisierung mentaler Modelle	24
Abbildung 6: Proposition P bestehend aus zwei Knoten K und einer Relation R	40
Abbildung 7: Modelle mit verschiedenen Maßen der Zerklüftetheit Z	42
Abbildung 8: Beispiel eines Modells und dessen aufbereitete Listenform zum Messzeitpunkt 1	43
Abbildung 9: Die drei Ebenen der SMD-Technologie	45
Abbildung 10: Beispiel-MODELL	48
Abbildung 11: Die einzelnen Schritte der SMD-Technologie	51
Abbildung 12: Rerepräsentation eines MODELLS der VPN 207 zum MZP 1	52
Abbildung 13: WEB-Interface der automatisierten SMD-Technologie	54
Abbildung 14: Ergebnisausgabe der automatisierten SMD-Technologie	55
Abbildung 15: Allgemeine experimentelle Anordnung der Veränderungsmessung in Anlehnung an Kleiter (1987)	59
Abbildung 16: Der Regressionseffekt am Beispiel eines Wissenstest (vgl. Petermann, 1978, p. 29)	65
Abbildung 17: Allgemeines Design der multimedialen Lernumgebung	82
Abbildung 18: Allgemeines Versuchsdesign der drei experimentellen Studien	84
Abbildung 19: 2x2-faktorielles Design der experimentellen Studien	85
Abbildung 20: Untersuchungsplan der Studie 1	87
Abbildung 21: Untersuchungsplan der Studie 2	89
Abbildung 22: Untersuchungsplan der Studie 3	91
Abbildung 23: Oberflächen-Struktur-Vergleich der ersten experimentellen Studie	111
Abbildung 24: Surface-Structure Index der ersten experimentellen Studie	113

Abbildung 25: Matching-Structure Index der ersten experimentellen Studie.....	114
Abbildung 26: Deep-Structure Index (Erklärungs- und Expertenmodell) der ersten experimentellen Studie	114
Abbildung 27: Oberflächen-Struktur-Vergleich der ersten experimentellen Studie getrennt nach den Untersuchungsgruppen.....	115
Abbildung 28: Surface-Structure Index der ersten experimentellen Studie getrennt nach den Untersuchungsgruppen.....	117
Abbildung 29: Matching-Structure Index der ersten experimentellen Studie getrennt nach den Untersuchungsgruppen.....	117
Abbildung 30: Deep-Structure Index der ersten experimentellen Studie getrennt nach den Untersuchungsgruppen und Ähnlichkeit zum Erklärungs- bzw. Expertenmodell	118
Abbildung 31: Oberflächen-Struktur-Vergleich der zweiten experimentellen Studie	129
Abbildung 32: Surface-Structure Index der zweiten experimentellen Studie.....	130
Abbildung 33: Matching-Structure Index der zweiten experimentellen Studie.....	131
Abbildung 34: Deep-Structure Index (Erklärungs- und Expertenmodell) der zweiten experimentellen Studie	132
Abbildung 35: Oberflächen-Struktur-Vergleich der zweiten experimentellen Studie getrennt nach den Untersuchungsgruppen.....	133
Abbildung 36: Surface-Structure Index der zweiten experimentellen Studie getrennt nach den Untersuchungsgruppen.....	134
Abbildung 37: Matching-Structure Index der zweiten experimentellen Studie getrennt nach den Untersuchungsgruppen.....	135
Abbildung 38: Deep-Structure Index der zweiten experimentellen Studie getrennt nach den Untersuchungsgruppen und Ähnlichkeit zum Erklärungs- bzw. Expertenmodell	136
Abbildung 39: Oberflächen-Struktur-Vergleich der dritten experimentellen Studie ..	146
Abbildung 40: Surface-Structure Index der dritten experimentellen Studie	147
Abbildung 41: Matching-Structure Index der dritten experimentellen Studie.....	148
Abbildung 42: Deep-Structure Index (Erklärungs- und Expertenmodell) der dritten experimentellen Studie	149
Abbildung 43: Oberflächen-Struktur-Vergleich der dritten experimentellen Studie getrennt nach den Untersuchungsgruppen.....	149
Abbildung 44: Surface-Structure Index der dritten experimentellen Studie getrennt nach den Untersuchungsgruppen.....	151

Abbildung 45: Matching-Structure Index der dritten experimentellen Studie getrennt nach den Untersuchungsgruppen.....	152
Abbildung 46: Deep-Structure Index der dritten experimentellen Studie getrennt nach den Untersuchungsgruppen und Ähnlichkeit zum Erklärungs- bzw. Expertenmodell	153
Abbildung 47: Erklärungsmodell der ersten experimentellen Studie (Geologie)	204
Abbildung 48: Expertenmodell der ersten experimentellen Studie (Geologie)	204
Abbildung 49: Erklärungsmodell der 2. & 3. experimentellen Studie (Geophysik)....	206
Abbildung 50: Expertenmodell der 2. & 3. experimentellen Studie (Geophysik).....	206

TABELLENVERZEICHNIS

Tabelle 1: Reliabilität der manuellen Analyse der SMD-Technologie.....	53
Tabelle 2: Verteilung der Probanden zu den experimentellen Bedingungen (Studie 1)	95
Tabelle 3: Verteilung der Probanden zu den experimentellen Bedingungen (Studie 2)	95
Tabelle 4: Verteilung der Probanden zu den experimentellen Bedingungen (Studie 3)	96
Tabelle 5: Reliabilität der Transformation der MODELLE in Listenform.....	98
Tabelle 6: Vergleich der vier Methodologien zur Erfassung von externalisierten Modellen (Johnson et al., 2006)	98
Tabelle 7: Test-Retest-Reliabilität des Oberflächen-Struktur-Vergleichs	99
Tabelle 8: Überprüfung der internen Validität der experimentellen Studien nach Campbell & Stanley (1966)	100
Tabelle 9: Test-Retest-Reliabilität für die Surface-, Matching- und Deep-Structure der Kontrollgruppe (N = 10).....	103
Tabelle 10: Bivariate Korrelationen nach Pearson zwischen Analyseergebnissen von MITOCAR und der SMD-Technologie für die experimentelle Studie 3 (N = 47).....	104
Tabelle 11: Bivariate Korrelationen nach Pearson bzw. Spearman zwischen der SMD- Technologie und dem deklarativen Wissen für die experimentelle Studie 3 (N = 47)	105
Tabelle 12: Effektstärkemaße und deren Interpretationsrichtwerte (Rost, 2005, p. 173)	107
Tabelle 13: Absolute Häufigkeiten und Prozentangaben der Untersuchungsstichprobe für die experimentellen Bedingungen <i>self-guided learning</i> (N = 13) und <i>scaffolding-based learning</i> (N = 13), Ergebnisse des Chi-Quadrat-Tests (Studie 1).....	110
Tabelle 14: Mittelwerte und Standardabweichungen der Untersuchungsstichprobe für die experimentellen Bedingungen <i>self-guided learning</i> (N = 13) und <i>scaffolding-based learning</i> (N = 13), Ergebnisse des Mann-Whitney-U-Tests (Studie 1).....	110
Tabelle 15: Durchschnittliche Veränderungswahrscheinlichkeit der MODELLE (N = 26) nach dem Oberflächen-Struktur-Vergleich (Studie 1).....	112
Tabelle 16: Ergebnisse der SMD-Technologie für die Messzeitpunkte der ersten experimentellen Studie (N = 26)	112

Tabelle 17: Durchschnittliche Veränderungswahrscheinlichkeit der MODELLE getrennt nach den Untersuchungsgruppen (Oberflächen-Struktur-Vergleich; Studie 1).....	116
Tabelle 18: Skalenwerte für deklaratives Wissen, Lernstrategien und allgemeine kognitive Fähigkeiten getrennt nach den Untersuchungsgruppen (Studie 1)	119
Tabelle 19: Lineares Wachstumsmodell der Surface-Structure (Studie 1)	121
Tabelle 20: Lineares Wachstumsmodell der Matching-Structure (Studie 1)	122
Tabelle 21: Lineares Wachstumsmodell der Deep-Structure, Erklärungsmodell (Studie 1)	122
Tabelle 22: Lineares Wachstumsmodell der Deep-Structure, Expertenmodell (Studie 1)	123
Tabelle 23: Absolute Häufigkeiten und Prozentangaben der Untersuchungsstichprobe für die experimentellen Bedingungen <i>self-guided learning</i> (N = 16) und <i>scaffolding-based learning</i> (N = 17), Ergebnisse des Chi-Quadrat-Tests (Studie 2).....	127
Tabelle 24: Mittelwerte und Standardabweichungen der Untersuchungsstichprobe für die experimentellen Bedingungen <i>self-guided learning</i> (N = 16) und <i>scaffolding-based learning</i> (N = 17), Ergebnisse des Mann-Whitney-U-Tests (Studie 2).....	128
Tabelle 25: Mittelwerte und Standardabweichungen der Untersuchungsstichprobe für die experimentellen Bedingungen <i>self-guided learning</i> (N = 16) und <i>scaffolding-based learning</i> (N = 17), Ergebnisse des t-Tests (Studie 2)	128
Tabelle 26: Durchschnittliche Veränderungswahrscheinlichkeit der MODELLE (N = 33) nach dem Oberflächen-Struktur-Vergleich (Studie 2).....	129
Tabelle 27: Ergebnisse der SMD-Technologie für die Messzeitpunkte der zweiten experimentellen Studie (N = 33)	130
Tabelle 28: Durchschnittliche Veränderungswahrscheinlichkeit der MODELLE getrennt nach den Untersuchungsgruppen (Oberflächen-Struktur-Vergleich; Studie 2).....	132
Tabelle 29: Skalenwerte für deklaratives Wissen, Lernstrategien und allgemeine kognitive Fähigkeiten getrennt nach den Untersuchungsgruppen (Studie 2)	137
Tabelle 30: Lineares Wachstumsmodell der Surface-Structure (Studie 2)	138
Tabelle 31: Lineares Wachstumsmodell der Matching-Structure (Studie 2)	139
Tabelle 32: Lineares Wachstumsmodell der Deep-Structure, Erklärungsmodell (Studie 2)	140

Tabelle 33: Lineares Wachstumsmodell der Deep-Structure, Expertenmodell (Studie 2)	140
Tabelle 34: Absolute Häufigkeiten und Prozentangaben der Untersuchungsstichprobe für die experimentellen Bedingungen <i>self-guided learning</i> (N = 18), <i>scaffolding-based learning</i> (N = 19) und <i>Kontrollgruppe</i> (N = 10), Ergebnisse des Chi-Quadrat-Tests (Studie 3)	144
Tabelle 35: Mittelwerte und Standardabweichungen der Untersuchungsstichprobe für die experimentellen Bedingungen <i>self-guided learning</i> (N = 18), <i>scaffolding-based learning</i> (N = 19) und <i>Kontrollgruppe</i> (N = 10), Ergebnisse des Kruskal-Wallis-H-Tests (Studie 3)	145
Tabelle 36: Mittelwerte und Standardabweichungen der Untersuchungsstichprobe für die experimentellen Bedingungen <i>self-guided learning</i> (N = 18), <i>scaffolding-based learning</i> (N = 19) und <i>Kontrollgruppe</i> (N = 10), Ergebnisse der ANOVA (Studie 3)	145
Tabelle 37: Durchschnittliche Veränderungswahrscheinlichkeit der MODELLE (N = 47) nach dem Oberflächen-Struktur-Vergleich (Studie 3)	146
Tabelle 38: Ergebnisse der SMD-Technologie für die Messzeitpunkte der dritten experimentellen Studie (N = 47)	147
Tabelle 39: Durchschnittliche Veränderungswahrscheinlichkeit der MODELLE getrennt nach den Untersuchungsgruppen (Oberflächen-Struktur-Vergleich; Studie 3)	150
Tabelle 40: Skalenwerte für deklaratives Wissen, Lernstrategien und allgemeine kognitive Fähigkeiten getrennt nach den Untersuchungsgruppen (Studie 3)	154
Tabelle 41: Lineares Wachstumsmodell der Surface-Structure (Studie 3)	155
Tabelle 42: Lineares Wachstumsmodell der Matching-Structure (Studie 3)	156
Tabelle 43: Lineares Wachstumsmodell der Deep-Structure, Erklärungsmodell (Studie 3)	156
Tabelle 44: Lineares Wachstumsmodell der Deep-Structure, Expertenmodell (Studie 3)	157
Tabelle 45: Ergebnisse der Surface-Structure nach Untersuchungsgruppen (Studie 1)	208
Tabelle 46: Ergebnisse der Matching-Structure nach Untersuchungsgruppen (Studie 1)	208
Tabelle 47: Ergebnisse der Deep-Structure nach Untersuchungsgruppen (Studie 1)	208
Tabelle 48: HLM-Analysen der Surface-Structure, getrennt nach den Prädiktoren GRUPPE und VORWISSEN (Studie 1)	209

Tabelle 49: HLM-Analysen der Matching-Structure, getrennt nach den Prädiktoren <i>GRUPPE</i> und <i>VORWISSEN</i> (Studie 1)	210
Tabelle 50: HLM-Analysen der Deep-Structure (Erklärungsmodell), getrennt nach den Prädiktoren <i>GRUPPE</i> und <i>VORWISSEN</i> (Studie 1)	211
Tabelle 51: HLM-Analysen der Deep-Structure (Expertenmodell), getrennt nach den Prädiktoren <i>GRUPPE</i> und <i>VORWISSEN</i> (Studie 1)	212
Tabelle 52: Ergebnisse der Surface-Structure nach Untersuchungsgruppen (Studie 2)	213
Tabelle 53: Ergebnisse der Matching-Structure nach Untersuchungsgruppen (Studie 2)	213
Tabelle 54: Ergebnisse der Deep-Structure nach Untersuchungsgruppen (Studie 2)	213
Tabelle 55: HLM-Analysen der Surface-Structure, getrennt nach den Prädiktoren <i>GRUPPE</i> und <i>VORWISSEN</i> (Studie 2)	214
Tabelle 56: HLM-Analysen der Matching-Structure, getrennt nach den Prädiktoren <i>GRUPPE</i> und <i>VORWISSEN</i> (Studie 2)	215
Tabelle 57: HLM-Analysen der Deep-Structure (Erklärungsmodell), getrennt nach den Prädiktoren <i>GRUPPE</i> und <i>VORWISSEN</i> (Studie 2)	216
Tabelle 58: HLM-Analysen der Deep-Structure (Expertenmodell), getrennt nach den Prädiktoren <i>GRUPPE</i> und <i>VORWISSEN</i> (Studie 2)	217
Tabelle 59: Ergebnisse der Surface-Structure nach Untersuchungsgruppen (Studie 3)	218
Tabelle 60: Ergebnisse der Matching-Structure nach Untersuchungsgruppen (Studie 3)	218
Tabelle 61: Ergebnisse der Deep-Structure (Erklärungsmodell) nach Untersuchungsgruppen (Studie 3)	219
Tabelle 62: Ergebnisse der Deep-Structure (Expertenmodell) nach Untersuchungsgruppen (Studie 3)	219
Tabelle 63: HLM-Analysen der Surface-Structure, getrennt nach den Prädiktoren <i>GRUPPE</i> und <i>VORWISSEN</i> (Studie 3)	220
Tabelle 64: HLM-Analysen der Matching-Structure, getrennt nach den Prädiktoren <i>GRUPPE</i> und <i>VORWISSEN</i> (Studie 3)	221
Tabelle 65: HLM-Analysen der Deep-Structure (Erklärungsmodell), getrennt nach den Prädiktoren <i>GRUPPE</i> und <i>VORWISSEN</i> (Studie 3)	222
Tabelle 66: HLM-Analysen der Deep-Structure (Expertenmodell), getrennt nach den Prädiktoren <i>GRUPPE</i> und <i>VORWISSEN</i> (Studie 3)	223

A Fragebogen zum Oberflächen-Struktur-Vergleich (Studie 3; Auszug)

Wissensdiagnose und mentale Modelle



Institut für Erziehungswissenschaft

VPN-Nummer _____
 Alter _____ Jahre
 Studiengang _____
 Hochschulse semester _____

- +** ähnliche Struktur
- verschiedene Struktur

Modell-Nr.	vgl. 1	vgl. 2	vgl. 3	vgl. 4	vgl. 5	vgl. 6
300						
301						
302						
303						
304						
305						
306						
307						
308						
309						
310						
311						
313						
314						
315						
316						

Vielen Dank.

B Erklärungs- und Expertenmodell⁴⁵ der ersten experimentellen Studie (Inhaltsdomäne Geologie)

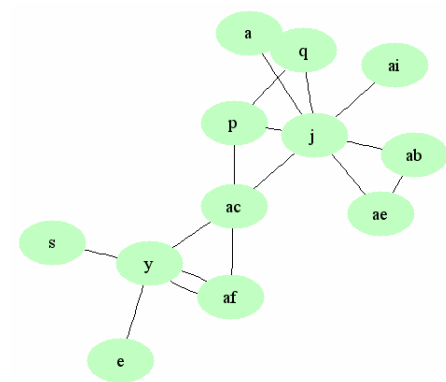


Abbildung 47: Erklärungsmodell der ersten experimentellen Studie (Geologie)

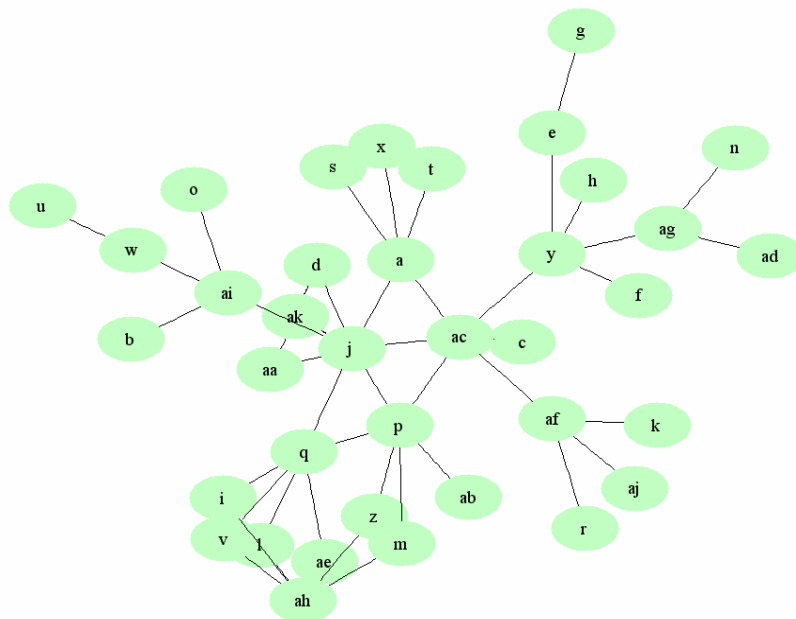


Abbildung 48: Expertenmodell der ersten experimentellen Studie (Geologie)

⁴⁵ Aus Übersichtsgründen werden die Modelle ohne benannte Relationen dargestellt.

C Lexikon zur Aufbereitung der MODELLE in Listenform (erste experimentelle Studie; Geologie)

Kürzel	semantische Zuordnung	Kürzel	semantische Zuordnung
A	anorganisches Material	AI	Verwesungsprodukte
B	Ausscheidung Zersetzung	AJ	Wärme
C	autotropes Lebewesen	AK	Zersetzung
D	Bakterien	AL	Nahrungskette
E	Baum	AM	Überflutung
F	Blume	AN	Lebensraum
G	Buche	AO	Schutz
H	Busch	AP	Eruption
I	Carnivoren	AQ	Klima
J	Destruent	AR	Lärm
K	Energie	AS	Gattung
L	Fleischfresser	AT	Ozon
M	Herivoren	AU	Biomasse
N	Himbeere	AV	Eigenschaft
O	Kohlenstoffverbindung	AW	Nähstoff
P	Konsument 1.Ordnung	AX	Chlorophyll
Q	Konsument 2. Ordnung	AY	Regen
R	Licht	ZE	Abholzung
S	Mineralstoffe	AZ	Darm
T	Nährsalze	BA	Organismus
U	nährstoffreicher Boden	BB	Erholung
V	Omnivoren	BC	Schaden
W	organisches Material	BD	Population
X	Oxide	BE	Oberfläche
Y	Pflanze	BF	belebt
Z	Pflanzenfresser	BG	Monokultur
AA	Pilze	BH	Schädling
AB	Primärkonsument	BI	Leben
AC	Produzent	BJ	Lebewesen
AD	Rose	BK	Verdunstung
AE	Sekundärkonsument	BL	Wolkenbildung
AF	Sonne	BM	Samen
AG	Strauch	BN	Wurzeln
AH	Tier		

D Erklärungs- und Expertenmodell der 2. & 3. experimentellen Studie (Inhaltsdomäne Geophysik)

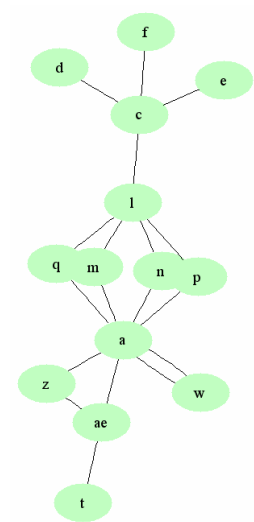


Abbildung 49: Erklärungsmodell der 2. & 3. experimentellen Studie (Geophysik)

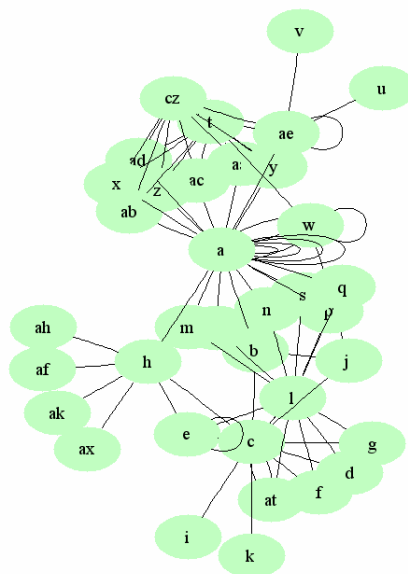


Abbildung 50: Expertenmodell der 2. & 3. experimentellen Studie (Geophysik)

E Lexikon zur Aufbereitung der MODELLE in Listenform (2. & 3. experimentelle Studie; Geophysik)

Kürzel	semantische Zuordnung	Kürzel	semantische Zuordnung
A	Schwarzwald	C	exogene Größe
AA	Faltung	CA	hart
AB	Verfestigung	CB	dick
AC	Anatexis	CC	Verschmutzung
AD	Metamorphose	CD	Theorie
AE	Plattentektonik	CE	Aufteilung
AF	Tiere	CF	Druck
AG	Artenvielfalt	CG	Stadt
AH	Pflanzen	CH	Subduktion
AI	Erde	CI	Umwelt
AJ	Nährstoffe	CJ	Zeit
AK	Lebewesen	CK	Beschaffenheit
AL	Lithosphäre	CL	Ränder
AM	Gravitation	CM	relief
AN	Partikelbestrahlung	CN	abrutschen
AO	Lava	CO	Höhe
AP	Rhein	CP	schwächer
AQ	Schicht	CQ	stärker
AR	Wärme	CR	externe Faktoren
AS	Wachstum	CS	Start
AT	Eis	CT	Verwehung
AV	Frostsprengung	CU	Überlagerung
AW	Gestein	CV	Position
AX	Menschen	CW	Kräfteeinwirkung
AY	Tal	CX	Transformstörung
AZ	Meteoriten	CY	verstehen
B	Naturkatastrophe	cz	endogene Größe
BA	jetzt	D	Wetter
BB	Zukunft	E	Klima
BC	Küste	F	Wind
BD	Gefahr	G	Luftdruck
BE	Angriffe aus dem All	H	Biosphäre
BF	Wüste	I	Atmosphäre
BG	Ernte	J	Ozeane
BH	Arbeitslose	K	Geomorphologie
BJ	Auswandern	L	exogener Prozess
BK	Überbevölkerung	M	Verwitterung
BL	Neulandbildung	N	Erosion
BM	Aussterben	O	Einebnung
BN	Atolle	P	Abtragung
BO	Industrie	Q	Transport
BP	Schlaglöcher	R	Ablagerung
BQ	Umsatz	S	Sedimentation
BR	Sauerstoff	T	endogener Prozess
BS	Teilchen	U	Vulkanismus
BT	Blick	V	Erdbeben
BU	Windrad	W	Formveränderung
BV	Strom	X	Umformung
BW	Sonne	Y	Hebung
BX	Fruchtbarkeit	Z	Senkung
BY	Naturgebiet	ZAH	Abholzung
BZ	Oberfläche		

F Detaillierte Ergebnisdarstellung der ersten Studie

Tabelle 45: Ergebnisse der Surface-Structure nach Untersuchungsgruppen (Studie 1)

Messzeit- punkt	Surface-Structure			
	scaffolding-based learning		self-guided learning	
	M	SD	M	SD
MZP 1	5,1538	3,6480	5,6154	3,1235
MZP 2	6,3846	6,8985	10,3846	8,9400
MZP 3	16,9231	6,1436	22,4615	15,8882
MZP 4	18,1538	5,8144	23,6923	16,7102
MZP 5	19,7692	7,8544	23,6923	16,7102
MZP 6	19,3846	6,5133	23,5385	16,7562
MZP 7	9,8462	4,4318	13,2308	6,5467

Tabelle 46: Ergebnisse der Matching-Structure nach Untersuchungsgruppen (Studie 1)

Messzeit- punkt	Matching-Structure			
	scaffolding-based learning		self-guided learning	
	M	SD	M	SD
MZP 1	1,6923	0,6304	2,3846	1,3868
MZP 2	1,1538	0,9871	2,1538	2,1153
MZP 3	2,5385	0,9674	3,5385	1,5607
MZP 4	2,6154	0,9608	3,6154	1,6093
MZP 5	2,8462	0,8987	3,6154	1,6093
MZP 6	2,6923	0,9473	3,6154	1,6093
MZP 7	2,6154	0,7679	3,0769	0,6405

Tabelle 47: Ergebnisse der Deep-Structure nach Untersuchungsgruppen (Studie 1)

Messzeit- punkt	Deep-Structure Erklärungsmodell				Deep-Structure Expertenmodell			
	scaffolding-based learning		self-guided learning		scaffolding-based learning		self-guided learning	
	M	SD	M	SD	M	SD	M	SD
MZP 1	0,0484	0,0695	0,0565	0,0734	0,0209	0,0301	0,0148	0,0248
MZP 2	0,0314	0,0639	0,0727	0,0771	0,0104	0,0209	0,0224	0,0335
MZP 3	0,1007	0,0985	0,0931	0,0841	0,0382	0,0378	0,0512	0,0491
MZP 4	0,1138	0,1226	0,0991	0,0856	0,0476	0,0640	0,0566	0,0547
MZP 5	0,1264	0,1010	0,0991	0,0856	0,0453	0,0576	0,0566	0,0547
MZP 6	0,1286	0,1055	0,0998	0,0863	0,0463	0,0605	0,0544	0,0523
MZP 7	0,1030	0,1023	0,1384	0,1017	0,0277	0,0334	0,0460	0,0491

Tabelle 48: HLM-Analysen der Surface-Structure, getrennt nach den Prädiktoren *GRUPPE* und *VORWISSEN* (Studie 1)

Prädiktoren	Surface-Structure				
	Fixe Effekte	Coefficient	SE	t-Wert (df=24)	p
<i>GRUPPE</i>	Mean Initial Status π_{0i}	15.175	4.831	3.141	.005
	<i>GRUPPE</i>	-3.057	3.055	-1.001	.327
	Mean Growth Rate π_{1i}	2.065	1.200	1.720	.098
	<i>GRUPPE</i>	-0.266	0.759	-0.351	.728
	Zufällige Effekte	Variance Component	SD	χ^2 (df=24)	p
	Initial Status r_{0i}	18.317	4.279	18.806	>.500
<i>VORWISSEN</i>	Growth Rate r_{1i}	0.489	0.699	10.487	>.500
	Level-1 Error e_{ti}	91.282	9.554		
	Fixe Effekte	Coefficient	SE	t-Wert (df=24)	p
	Mean Initial Status π_{0i}	10.895	2.200	4.951	.000
	<i>VoWi</i>	-0.612	3.112	-0.197	.846
	Mean Growth Rate π_{1i}	1.708	0.538	3.174	.005
<i>VORWISSEN</i>	<i>VoWi</i>	-0.085	0.761	-0.112	.912
	Zufällige Effekte	Variance Component	SD	χ^2 (df=24)	p
	Initial Status r_{0i}	20.561	4.534	20.177	>.500
	Growth Rate r_{1i}	0.506	0.711	10.612	>.500
	Level-1 Error e_{ti}	91.307	9.555		

Tabelle 49: HLM-Analysen der Matching-Structure, getrennt nach den Prädiktoren *GRUPPE* und *VORWISSEN* (Studie 1)

Prädiktoren	Matching-Structure				
	Fixe Effekte	Coefficient	SE	t-Wert (df=24)	p
<i>GRUPPE</i>	Mean Initial Status π_{0i}	3.549	0.577	6.145	<.001
	<i>GRUPPE</i>	-0.950	0.365	-2.602	.016
	Mean Growth Rate π_{1i}	0.142	0.130	1.098	.284
	<i>GRUPPE</i>	0.038	0.082	0.467	.644
	Zufällige Effekte	Variance Component	SD	χ^2 (df=24)	p
	Initial Status r_{0i}	0.345	0.587	33.537	.093
<i>VORWISSEN</i>	Growth Rate r_{1i}	0.003	0.062	18.832	>.500
	Level-1 Error e_{ti}	1.123	1.060		
	Fixe Effekte	Coefficient	SE	t-Wert (df=24)	p
	Mean Initial Status π_{0i}	2.197	0.288	7.626	<.001
	<i>VoWi</i>	-0.148	0.407	-0.364	.719
	Mean Growth Rate π_{1i}	0.197	0.057	3.450	.002
<i>VORWISSEN</i>	<i>VoWi</i>	0.005	0.081	0.068	.947
	Zufällige Effekte	Variance Component	SD	χ^2 (df=24)	p
	Initial Status r_{0i}	0.553	0.743	44.134	.008
	Growth Rate r_{1i}	0.002	0.047	18.901	>.500
	Level-1 Error e_{ti}	1.133	1.064		

Tabelle 50: HLM-Analysen der Deep-Structure (Erklärungsmodell), getrennt nach den Prädiktoren *GRUPPE* und *VORWISSEN* (Studie 1)

Prädiktoren	Deep-Structure (Erklärungsmodell)				
	Fixe Effekte	Coefficient	SE	t-Wert (df=24)	p
<i>GRUPPE</i>	Mean Initial Status π_{0i}	0.057	0.033	1.716	.099
	<i>GRUPPE</i>	-0.002	0.021	-0.122	.904
	Mean Growth Rate π_{1i}	0.010	0.009	1.115	.276
	<i>GRUPPE</i>	0.001	0.006	0.237	.815
	Zufällige Effekte	Variance Component	SD	χ^2 (df=24)	p
	Initial Status r_{0i}	0.00113	0.033	38.710	.029
	Growth Rate r_{1i}	0.00011	0.010	42.617	.011
	Level-1 Error e_{ti}	0.00378	0.061		
<i>VORWISSEN</i>	Fixe Effekte	Coefficient	SE	t-Wert (df=24)	p
	Mean Initial Status π_{0i}	0.042	0.014	2.881	.009
	<i>VoWi</i>	0.021	0.020	1.055	.303
	Mean Growth Rate π_{1i}	0.014	0.004	3.319	.003
	<i>VoWi</i>	-0.002	0.006	-0.473	.640
	Zufällige Effekte	Variance Component	SD	χ^2 (df=24)	p
	Initial Status r_{0i}	0.00106	0.032	37.171	.042
	Growth Rate r_{1i}	0.00011	0.010	42.558	.011
	Level-1 Error e_{ti}	0.00376	0.061		

Tabelle 51: HLM-Analysen der Deep-Structure (Expertenmodell), getrennt nach den Prädiktoren *GRUPPE* und *VORWISSEN* (Studie 1)

Prädiktoren	Deep-Structure (Erxpertenmodell)				
	Fixe Effekte	Coefficient	SE	t-Wert (df=24)	p
<i>GRUPPE</i>	Mean Initial Status π_{0i}	0.028	0.015	1.865	.074
	<i>GRUPPE</i>	-0.002	0.009	-0.266	.792
	Mean Growth Rate π_{1i}	0.008	0.005	1.496	.148
	<i>GRUPPE</i>	-0.002	0.003	-0.665	.512
	Zufällige Effekte	Variance Component	SD	χ^2 (df=24)	p
	Initial Status r_{0i}	0.00021	0.014	31.770	.133
<i>VORWISSEN</i>	Growth Rate r_{1i}	0.00005	0.006	57.195	<.001
	Level-1 Error e_{ti}	0.00082	0.028		
	Fixe Effekte	Coefficient	SE	t-Wert (df=24)	p
	Mean Initial Status π_{0i}	0.022	0.006	3.380	.003
	<i>VoWi</i>	0.003	0.009	0.320	.751
	Mean Growth Rate π_{1i}	0.004	0.002	1.855	.076
<i>VORWISSEN</i>	<i>VoWi</i>	0.0003	0.0034	0.090	.929
	Zufällige Effekte	Variance Component	SD	χ^2 (df=24)	p
	Initial Status r_{0i}	0.00021	0.014	31.721	.134
	Growth Rate r_{1i}	0.00005	0.006	58.325	<.001
	Level-1 Error e_{ti}	0.00082	0.028		

G Detaillierte Ergebnisdarstellung der zweiten Studie

Tabelle 52: Ergebnisse der Surface-Structure nach Untersuchungsgruppen (Studie 2)

Messzeit- punkt	Surface-Structure			
	scaffolding-based learning		self-guided learning	
	M	SD	M	SD
MZP 1	5,7647	1,7864	5,8750	3,7036
MZP 2	1,7647	2,6346	6,2500	5,3104
MZP 3	8,0000	4,6098	8,6875	5,6771
MZP 4	11,8824	6,1021	9,8125	5,6002
MZP 5	12,4706	6,9202	9,4375	5,1117
MZP 6	12,9412	6,9953	9,6875	4,9493
MZP 7	9,6471	4,8469	7,2500	2,4358

Tabelle 53: Ergebnisse der Matching-Structure nach Untersuchungsgruppen (Studie 2)

Messzeit- punkt	Matching-Structure			
	scaffolding-based learning		self-guided learning	
	M	SD	M	SD
MZP 1	2,5882	0,7952	2,7500	0,8563
MZP 2	1,0000	1,5000	2,2500	1,3904
MZP 3	3,1176	1,6912	2,9375	1,4361
MZP 4	4,0000	1,8028	3,6250	1,7078
MZP 5	4,2353	1,9212	3,6250	1,7464
MZP 6	4,1765	1,8109	3,6875	1,8154
MZP 7	3,8235	1,3800	3,3125	1,7017

Tabelle 54: Ergebnisse der Deep-Structure nach Untersuchungsgruppen (Studie 2)

Messzeit- punkt	Deep-Structure Erklärungsmodell				Deep-Structure Expertenmodell			
	scaffolding-based learning		self-guided learning		scaffolding-based learning		self-guided learning	
	M	SD	M	SD	M	SD	M	SD
MZP 1	0,0241	0,0568	0,0500	0,0664	0,0271	0,0614	0,0262	0,0289
MZP 2	0,0145	0,0437	0,1212	0,1238	0,0144	0,0299	0,0736	0,0872
MZP 3	0,1039	0,1119	0,1347	0,1238	0,0816	0,0997	0,0836	0,0699
MZP 4	0,1180	0,1249	0,1152	0,1068	0,1099	0,1052	0,0767	0,0707
MZP 5	0,1454	0,1443	0,1169	0,0987	0,1175	0,1074	0,0693	0,0676
MZP 6	0,1379	0,1375	0,1225	0,1058	0,1180	0,1115	0,0675	0,0635
MZP 7	0,1090	0,1309	0,1699	0,1407	0,0842	0,1000	0,1111	0,0728

Tabelle 55: HLM-Analysen der Surface-Structure, getrennt nach den Prädiktoren *GRUPPE* und *VORWISSEN* (Studie 2)

Prädiktoren	Surface-Structure				
	Fixe Effekte	<i>Coefficient</i>	<i>SE</i>	<i>t-Wert</i> (<i>df</i> =31)	<i>p</i>
<i>GRUPPE</i>	Mean Initial Status π_{0i}	6.883	0.816	8.430	<.001
	<i>GRUPPE</i>	-2.081	1.137	-1.829	.076
	Mean Growth Rate π_{1i}	0.419	0.232	1.802	.081
	<i>GRUPPE</i>	0.954	0.324	2.941	.007
	Zufällige Effekte	<i>Variance Component</i>	<i>SD</i>	χ^2 (<i>df</i> =31)	<i>p</i>
	Initial Status r_{0i}	3.458	1.859	44.510	.055
	Growth Rate r_{1i}	0.313	0.559	47.178	.031
	Level-1 Error e_{ti}	15.528	3.940		
	Fixe Effekte	<i>Coefficient</i>	<i>SE</i>	<i>t-Wert</i> (<i>df</i> =31)	<i>p</i>
	Mean Initial Status π_{0i}	5.996	1.071	5.595	<.001
<i>VORWISSEN</i>	<i>VoWi</i>	-0.265	1.283	-0.206	.838
	Mean Growth Rate π_{1i}	0.525	0.318	1.646	.109
	<i>VoWi</i>	0.554	0.382	1.451	.157
	Zufällige Effekte	<i>Variance Component</i>	<i>SD</i>	χ^2 (<i>df</i> =31)	<i>p</i>
	Initial Status r_{0i}	4.215	2.053	48.969	.021
	Growth Rate r_{1i}	0.457	0.676	56.363	.004
	Level-1 Error e_{ti}	15.663	3.957		

Tabelle 56: HLM-Analysen der Matching-Structure, getrennt nach den Prädiktoren *GRUPPE* und *VORWISSEN* (Studie 2)

Prädiktoren	Matching-Structure				
	Fixe Effekte	Coefficient	SE	t-Wert (df=31)	p
GRUPPE	Mean Initial Status π_{0i}	2.607	0.240	10.855	<.001
	GRUPPE	-0.527	0.334	-1.576	.125
	Mean Growth Rate π_{1i}	0.187	0.073	2.546	.016
	GRUPPE	0.211	0.102	2.063	.047
	Zufällige Effekte	Variance Component	SD	χ^2 (df=31)	p
	Initial Status r_{0i}	0.294	0.542	42.082	.088
	Growth Rate r_{1i}	0.038	0.195	52.359	.010
	Level-1 Error e_{ti}	1.354	1.163		
	Fixe Effekte	Coefficient	SE	t-Wert (df=31)	p
	Mean Initial Status π_{0i}	2.546	0.301	8.456	<.001
VORWISSEN	VoWi	-0.302	0.360	-0.839	.408
	Mean Growth Rate π_{1i}	0.246	0.096	2.563	.016
	VoWi	0.071	0.115	0.624	.537
	Zufällige Effekte	Variance Component	SD	χ^2 (df=31)	p
	Initial Status r_{0i}	0.264	0.514	43.752	.064
	Growth Rate r_{1i}	0.043	0.207	57.977	.003
	Level-1 Error e_{ti}	1.383	1.176		
	Fixe Effekte	Coefficient	SE	t-Wert (df=31)	p
	Mean Initial Status π_{0i}	2.546	0.301	8.456	<.001
	VoWi	-0.302	0.360	-0.839	.408

Tabelle 57: HLM-Analysen der Deep-Structure (Erklärungsmodell), getrennt nach den Prädiktoren *GRUPPE* und *VORWISSEN* (Studie 2)

Prädiktoren		Deep-Structure (Erklärungsmodell)				
	Fixe Effekte	<i>Coefficient</i>	<i>SE</i>	<i>t-Wert</i> (<i>df</i> =31)	<i>p</i>	
<i>GRUPPE</i>	Mean Initial Status π_{0i}	0.081	0.014	5.471	<.001	
	<i>GRUPPE</i>	-0.046	0.020	-2.239	.032	
	Mean Growth Rate π_{1i}	0.0123	0.006	1.941	.061	
	<i>GRUPPE</i>	0.007	0.008	0.802	.429	
	Zufällige Effekte	<i>Variance Component</i>	<i>SD</i>	χ^2 (<i>df</i> =31)	<i>p</i>	
	Initial Status r_{0i}	0.00092	0.030	41.721	.095	
	Growth Rate r_{1i}	0.00044	0.020	97.664	<.001	
	Level-1 Error e_{ti}	0.00571	0.075			
<i>VORWISSEN</i>	Fixe Effekte	<i>Coefficient</i>	<i>SE</i>	<i>t-Wert</i> (<i>df</i> =31)	<i>p</i>	
	Mean Initial Status π_{0i}	0.044	0.020	2.190	.036	
	<i>VoWi</i>	0.019	0.024	0.807	.426	
	Mean Growth Rate π_{1i}	0.004	0.007	0.541	.592	
	<i>VoWi</i>	0.016	0.009	1.837	.075	
	Zufällige Effekte	<i>Variance Component</i>	<i>SD</i>	χ^2 (<i>df</i> =31)	<i>p</i>	
	Initial Status r_{0i}	0.00141	0.037	47.474	.029	
	Growth Rate r_{1i}	0.00039	0.019	89.911	<.001	
	Level-1 Error e_{ti}	0.00571	0.075			

Tabelle 58: HLM-Analysen der Deep-Structure (Expertenmodell), getrennt nach den Prädiktoren *GRUPPE* und *VORWISSEN* (Studie 2)

Prädiktoren	Deep-Structure (Erxpertenmodell)				
	Fixe Effekte	<i>Coefficient</i>	<i>SE</i>	<i>t-Wert</i> (<i>df</i> =31)	<i>p</i>
<i>GRUPPE</i>	Mean Initial Status π_{0i}	0.048	0.010	4.393	<.001
	<i>GRUPPE</i>	-0.013	0.015	-0.888	.381
	Mean Growth Rate π_{1i}	0.008	0.003	2.097	.044
	<i>GRUPPE</i>	0.006	0.005	1.226	.230
	Zufällige Effekte	<i>Variance Component</i>	<i>SD</i>	χ^2 (<i>df</i> =31)	<i>p</i>
	Initial Status r_{0i}	0.00069	0.026	48.224	.025
	Growth Rate r_{1i}	0.00015	0.012	79.184	<.001
	Level-1 Error e_{ti}	0.00266	0.051		
<i>VORWISSEN</i>	Fixe Effekte	<i>Coefficient</i>	<i>SE</i>	<i>t-Wert</i> (<i>df</i> =31)	<i>p</i>
	Mean Initial Status π_{0i}	0.016	0.013	1.292	.206
	<i>VoWi</i>	0.034	0.015	2.237	.033
	Mean Growth Rate π_{1i}	0.003	0.004	0.727	.472
	<i>VoWi</i>	0.011	0.005	2.065	.047
	Zufällige Effekte	<i>Variance Component</i>	<i>SD</i>	χ^2 (<i>df</i> =31)	<i>p</i>
	Initial Status r_{0i}	0.00046	0.021	42.580	.080
	Growth Rate r_{1i}	0.00013	0.011	72.983	<.001
	Level-1 Error e_{ti}	0.00266	0.051		

H Detaillierte Ergebnisdarstellung der dritten Studie

Tabelle 59: Ergebnisse der Surface-Structure nach Untersuchungsgruppen (Studie 3)

Messzeit- punkt	Surface-Structure					
	scaffolding-based learning		self-guided learning		Kontrollgruppe	
	M	SD	M	SD	M	SD
MZP 1	4,1111	3,0657	3,6111	3,1462	3,2000	2,7809
MZP 2	3,7222	3,9527	2,3889	3,6964	4,7000	3,8020
MZP 3	6,2222	5,4725	2,2778	2,6303	4,9000	3,7253
MZP 4	5,5556	4,6174	3,8889	3,0076	4,8000	4,1580
MZP 5	8,3333	4,3114	7,8889	6,7291	5,8000	4,6380
MZP 6	7,8889	5,6971	8,7778	7,8782	6,8000	2,7809
MZP 7	6,8889	4,2687	5,6667	4,9349	7,4000	4,3256

Tabelle 60: Ergebnisse der Matching-Structure nach Untersuchungsgruppen (Studie 3)

Messzeit- punkt	Matching-Structure					
	scaffolding-based learning		self-guided learning		Kontrollgruppe	
	M	SD	M	SD	M	SD
MZP 1	2,0556	1,2590	1,4444	1,1490	1,8000	1,4757
MZP 2	1,6667	1,6450	1,0556	1,3048	1,1000	1,1972
MZP 3	2,7222	2,0236	1,3333	1,5339	1,6000	1,1738
MZP 4	2,3333	1,6803	2,1111	1,5676	1,8000	1,2293
MZP 5	3,0000	1,4951	2,8889	1,8752	2,9000	1,7288
MZP 6	2,7778	1,5551	2,7222	1,2744	3,4000	1,5055
MZP 7	2,9444	1,5136	2,5000	1,5049	3,4000	1,5055

Tabelle 61: Ergebnisse der Deep-Structure (Erklärungsmodell) nach Untersuchungsgruppen (Studie 3)

Messzeitpunkt	Deep-Structure Erklärungsmodell					
	scaffolding-based learning		self-guided learning		Kontrollgruppe	
	M	SD	M	SD	M	SD
MZP 1	0,0708	0,1012	0,1039	0,1237	0,0461	0,0630
MZP 2	0,1057	0,1296	0,0504	0,1141	0,0910	0,1175
MZP 3	0,0924	0,1220	0,0973	0,1700	0,0752	0,1057
MZP 4	0,1032	0,1230	0,0725	0,0922	0,0830	0,0944
MZP 5	0,1731	0,1433	0,1043	0,1172	0,0424	0,0582
MZP 6	0,0993	0,0964	0,1055	0,1058	0,0766	0,0614
MZP 7	0,1056	0,1044	0,1301	0,1153	0,0951	0,1078

Tabelle 62: Ergebnisse der Deep-Structure (Expertenmodell) nach Untersuchungsgruppen (Studie 3)

Messzeitpunkt	Deep-Structure Expertenmodell					
	scaffolding-based learning		self-guided learning		Kontrollgruppe	
	M	SD	M	SD	M	SD
MZP 1	0,0771	0,0812	0,0665	0,0851	0,0654	0,1165
MZP 2	0,0569	0,0723	0,0288	0,0631	0,1164	0,1277
MZP 3	0,0581	0,0672	0,0434	0,0774	0,0754	0,0699
MZP 4	0,0694	0,0827	0,0422	0,0523	0,0561	0,0949
MZP 5	0,1140	0,1018	0,0866	0,0837	0,0260	0,0256
MZP 6	0,0710	0,0513	0,1005	0,1175	0,0429	0,0419
MZP 7	0,0657	0,0542	0,0792	0,0864	0,0416	0,0509

Tabelle 63: HLM-Analysen der Surface-Structure, getrennt nach den Prädiktoren *GRUPPE* und *VORWISSEN* (Studie 3)

Prädiktoren	Surface-Structure				
	Fixe Effekte	Coefficient	SE	t-Wert (df=45)	p
<i>GRUPPE</i>	Mean Initial Status π_{0i}	2.343	0.790	2.966	.005
	<i>GRUPPE</i>	0.809	0.558	1.448	.154
	Mean Growth Rate π_{1i}	0.508	0.265	1.918	.061
	<i>GRUPPE</i>	0.191	0.187	1.021	.313
	Zufällige Effekte	Variance Component	SD	χ^2 (df=45)	p
	Initial Status r_{0i}	1.690	1.300	56.140	.123
	Growth Rate r_{1i}	0.434	0.659	82.248	.001
	Level-1 Error e_{ti}	14.709	3.835		
	Fixe Effekte	Coefficient	SE	t-Wert (df=45)	p
	Mean Initial Status π_{0i}	3.066	0.620	4.942	<.001
<i>VORWISSEN</i>	<i>VoWi</i>	0.473	0.868	0.545	.588
	Mean Growth Rate π_{1i}	0.743	0.206	3.599	.001
	<i>VoWi</i>	0.289	-0.013	-0.045	.964
	Zufällige Effekte	Variance Component	SD	χ^2 (df=45)	p
	Initial Status r_{0i}	2.029	1.424	58.372	.087
	Growth Rate r_{1i}	0.457	0.676	84.149	.001
	Level-1 Error e_{ti}	14.709	3.835		
	Fixe Effekte	Coefficient	SE	t-Wert (df=45)	p
	Mean Initial Status π_{0i}	3.066	0.620	4.942	<.001
	<i>VoWi</i>	0.473	0.868	0.545	.588

Tabelle 64: HLM-Analysen der Matching-Structure, getrennt nach den Prädiktoren *GRUPPE* und *VORWISSEN* (Studie 3)

Prädiktoren		Matching-Structure			
	Fixe Effekte	<i>Coefficient</i>	<i>SE</i>	<i>t-Wert</i> (<i>df</i> =45)	<i>p</i>
<i>GRUPPE</i>	Mean Initial Status π_{0i}	1.010	0.290	3.471	.001
	<i>GRUPPE</i>	0.391	0.205	1.902	.063
	Mean Growth Rate π_{1i}	0.227	0.077	2.921	.006
	<i>GRUPPE</i>	0.035	0.055	0.642	.524
	Zufällige Effekte	<i>Variance Component</i>	<i>SD</i>	χ^2 (<i>df</i> =45)	<i>p</i>
	Initial Status r_{0i}	0.532	0.729	83.508	.001
	Growth Rate r_{1i}	0.034	0.186	77.831	.002
	Level-1 Error e_{ti}	1.340	1.157		
<i>VORWISSEN</i>	Fixe Effekte	<i>Coefficient</i>	<i>SE</i>	<i>t-Wert</i> (<i>df</i> =45)	<i>p</i>
	Mean Initial Status π_{0i}	1.453	0.232	6.242	<.001
	<i>VoWi</i>	0.045	0.325	0.138	.891
	Mean Growth Rate π_{1i}	0.059	0.298	4.971	<.001
	<i>VoWi</i>	-0.055	0.083	-0.662	.511
	Zufällige Effekte	<i>Variance Component</i>	<i>SD</i>	χ^2 (<i>df</i> =45)	<i>p</i>
	Initial Status r_{0i}	0.624	0.790	90.185	<.001
	Growth Rate r_{1i}	0.034	0.186	77.787	.002
	Level-1 Error e_{ti}	1.340	1.157		

Tabelle 65: HLM-Analysen der Deep-Structure (Erklärungsmodell), getrennt nach den Prädiktoren *GRUPPE* und *VORWISSEN* (Studie 3)

Prädiktoren	Deep-Structure (Erklärungsmodell)				
	Fixe Effekte	Coefficient	SE	t-Wert (df=45)	p
GRUPPE	Mean Initial Status π_{0i}	0.062	0.025	2.476	.017
	GRUPPE	0.013	0.017	0.754	.455
	Mean Growth Rate π_{1i}	0.001	0.006	0.223	.825
	GRUPPE	0.003	0.004	0.764	.449
	Zufällige Effekte	Variance Component	SD	χ^2 (df=45)	p
	Initial Status r_{0i}	0.00459	0.067	93.771	<.001
	Growth Rate r_{1i}	0.00026	0.016	81.328	.001
	Level-1 Error e_{ti}	0.00912	0.095		
VORWISSEN	Fixe Effekte	Coefficient	SE	t-Wert (df=45)	p
	Mean Initial Status π_{0i}	0.070	0.019	3.596	.001
	VoWi	0.016	0.027	0.608	.546
	Mean Growth Rate π_{1i}	0.005	0.005	1.023	.312
	VoWi	0.0009	0.007	0.134	.895
	Zufällige Effekte	Variance Component	SD	χ^2 (df=45)	p
	Initial Status r_{0i}	0.00463	0.068	94.182	<.001
	Growth Rate r_{1i}	0.00027	0.016	82.350	.001
	Level-1 Error e_{ti}	0.00912	0.095		

Tabelle 66: HLM-Analysen der Deep-Structure (Expertenmodell), getrennt nach den Prädiktoren *GRUPPE* und *VORWISSEN* (Studie 3)

Prädiktoren		Deep-Structure (Erxpertenmodell)				
	Fixe Effekte	<i>Coefficient</i>	<i>SE</i>	<i>t-Wert</i> (<i>df</i> =45)	<i>p</i>	
<i>GRUPPE</i>	Mean Initial Status π_{0i}	0.029	0.018	1.547	.129	
	<i>GRUPPE</i>	0.027	0.013	2.040	.047	
	Mean Growth Rate π_{1i}	0.004	0.005	0.848	.401	
	<i>GRUPPE</i>	-0.001	0.003	-0.501	.618	
	Zufällige Effekte	<i>Variance Component</i>	<i>SD</i>	χ^2 (<i>df</i> =45)	<i>p</i>	
	Initial Status r_{0i}	0.00304	0.055	120.018	<.001	
	Growth Rate r_{1i}	0.00022	0.014	114.830	<.001	
	Level-1 Error e_{ti}	0.00393	0.062			
<i>VORWISSEN</i>	Fixe Effekte	<i>Coefficient</i>	<i>SE</i>	<i>t-Wert</i> (<i>df</i> =45)	<i>p</i>	
	Mean Initial Status π_{0i}	0.045	0.014	3.056	.004	
	<i>VoWi</i>	0.032	0.020	1.550	.128	
	Mean Growth Rate π_{1i}	0.002	0.003	0.639	.526	
	<i>VoWi</i>	-0.0006	0.005	-0.123	.903	
	Zufällige Effekte	<i>Variance Component</i>	<i>SD</i>	χ^2 (<i>df</i> =45)	<i>p</i>	
	Initial Status r_{0i}	0.00322	0.056	124.473	<.001	
	Growth Rate r_{1i}	0.00022	0.014	115.432	<.001	
	Level-1 Error e_{ti}	0.00393	0.062			